SKRIPSI

PEMODELAN ANGKA PUTUS SEKOLAH BAGI ANAK USIA WAJIB BELAJAR DI PROVINSI SULAWESI SELATAN DENGAN PENDEKATAN GENERALIZED POISSON REGRESSION (GPR)

NASRA

1311141014

MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI MAKASSAR
2016/2017

SKRIPSI

PEMODELAN ANGKA PUTUS SEKOLAH BAGI ANAK USIA WAJIB BELAJAR DI PROVINSI SULAWESI SELATAN DENGAN PENDEKATAN GENERALIZED POISSON REGRESSION (GPR)

Diajukan kepada Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Makassar untuk memenuhi salah satu syarat memperoleh gelar Sarjana Sains Matematika

NASRA

1311141014

MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI MAKASSAR
2016/2017

HALAMAN PENGESAHAN

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini menyatakan bahwa skripsi ini adalah hasil karya saya sendiri, dan semua sumber baik yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar. Bila kemudian hari ternyata pernyataan saya terbukti tidak benar, maka saya bersedia menerima sanksi yang telah ditetapkan oleh FMIPA Universitas Negeri Makassar, Makassar.

Yang membuat pernyataan

Nama	: Nasra
NIM	: 1311141014
Tanggal	: Juli 2017

PERSETUJUAN PUBLIKASI

Sebagai civitas akademi Universitas Negeri Makassar, saya bertanda tangan di bawah ini :

Nama	: Nasra
NIM	: 1311141014
Program studi	: Matematika
Fakultas	: MIPA

demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Negeri Makassar **Hak Bebas Royalti Non-Ekslusif** (*Non-ekslusive Royalti Free Right*) atas skripsi yang berjudul; "Pemodelan Angka Putus Sekolah Bagi Anak Usia Wajib Belajar di Provinsi Sulawesi Selatan dengan Pendekatan *Generalized Poisson Regression* (GPR)", beserta perangkat yang ada (jika diperlukan). Dengan **Hak Bebas Royalti Non-Ekslusif**, Universitas Negeri Makassar berhak menyimpan, mengalih media/formatkan, mengelolah dalam bentuk pangkalan data (Data Base), merawat, dan mempublikasikan skripsi saya selama tetap mencamtumkan nama saya sebagai penulis, pencipta, dan pemilik hak cipta serta tidak dikomersilkan

Demikian pernyataan ini saya buat dengan sebenarnya

Dibuat di	: Makassar
Pada Tanggal	: Juli 2017

Menyetujui

Pembimbing I	Yang Menyatakan
Dr. Hisyam Ihsan, M.Si	**Nasra**
NIP. 19651226 199103 1 001	**NIM. 1311141014**

MOTTO DAN PERSEMBAHAN

“Boleh jadi kamu membenci sesuatu, padahaliaamat baik bagimu, dan boleh jadi kamu menyukai sesuatu padahal ia amat buruk bagimu, Allah mengetahui, sedang kamu tidak mengetahui” (QS. Al-Baqarah:216)

“Jika sore tiba, janganlah tunggu waktu pagi, jika pagi tiba, janganlah tunggu waktu sore. Manfaatkan masa sehatmu sebelum tiba masa sakitmu dan manfaatkan masa hidupmu sebelum tiba ajalmu” (Umar bin Khattab)

“Pamer adalah ide yang bodoh untuk sebuah kemenangan” (Bruce Lee)

“Anda mungkin bisa menunda tapi waktu tidak akan menunggu” (Benjamin Franklin)

Kupersembahkan karya sederhana ini untuk Ayahanda (Sudirman), Ibunda (Erna), Suami (Reski), Saudaraku (Nasri & Nasru), keluarga besarku dan orang-orang yang senantiasa mendoakanku dalam sujudnya serta semua pihak. Semoga karya sederhana ini bermanfaat. Aamiin yaa rabbal ‘alamin.

ABSTRACT

Nasra, 2017. Pemodelan Angka Putus Sekolah Bagi Anak Usia Wajib Belajar Di Provinsi Sulawesi Selatan Dengan Pendekatan Generalized Poisson Regression (GPR). Skripsi. Jurusan Matematika. Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Negeri Makassar. (Guided by Hisyam Ihsan dan. Wahidah Sanusi).

Dropout number is one of the important indicators to measure the human progress resources in education sector. In this research used the approaches of Generalized Poisson Regression to get the best model and to determine the influencing factors of droput number for children of compulsory age in Sulawesi Selatan. The result of this research show that factors which significantly influence the drop out number for children of compulsory age in Sulawesi Selatan are ratio of student to teacher, the number of poor people, population density, and the average lenght of school.

Kata kunci: *Generalized Poisson Regression*, *Dropout Number*

ABSTRAK

Nasra, 2017. Pemodelan Angka Putus Sekolah Bagi Anak Usia Wajib Belajar Di Provinsi Sulawesi Selatan Dengan Pendekatan Generalized Poisson Regression (GPR). Skripsi. Jurusan Matematika. Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Negeri Makassar. (Dibimbing oleh Hisyam Ihsan dan. Wahidah Sanusi)

Putus sekolah merupakan salah satu indikator yang berguna untuk mengukur kemajuan sumber daya manusia pada bidang pendidikan. Pada penitian ini digunakan pendekatan Generalized Poisson Regression untuk mendapatkan model terbaik dan faktor-faktor yang berpengaruh secara signifikan terhadap angka putus sekolah usia wajib belajar di Sulawesi Selatan. Hasil penelitian ini menujukkan bahwa faktor-faktor yang berpengaruh secara signifikan terhadap angka putus sekolah di Sulawesi Selatan adalah rasio siswa terhadap guru, jumlah penduduk miskin, kepadatan penduduk, dan rata-rata lama bersekolah.

Keywords: *Generalized Poisson Regression, Angka Putus Sekolah*

KATA PENGANTAR

Alhamdulillahi Rabbil ‘Alaamiin, segala puji hanya milik Allah *Subhanahu wa Ta’ala* pencipta alam semesta atas limpahan nikmat, selaras dengan penuh keagungan dan kebesaran-Nya sehingga penulis dapat menyelesaikan tugas akhir ini. Salam dan shalawat senantiasa terkirim untuk Baginda Nabi Besar Rasulullah Muhammad Shallallahu ‘alaihi wa sallam dan semoga keselamatan senantiasa tercurahkan kepada keluarga beliau, para sahabat, *tabi’in, tabi’ut-tabi’in*, serta umat yang senantiasa istiqomah di jalan-Nya.

Skripsi yang berjudul “Pemodelan Angka Putus Sekolah Bagi Anak Usia Wajib Belajar di Provinsi Sulawesi Selatan dengan Pendekatan *Generalized Poisson Regression*” ini disusun sebagai syarat untuk memperoleh gelar Sarjana Sains di bidang Matematika pada Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Negeri Makassar (UNM).

Penulis menyadari bahwa sejak proses penyusunan proposal hingga pembuatan laporan hasil penelitian yang dituangkan dalam skripsi ini tidak sedikit hambatan dan tantangan yang dihadapi penulis. Namun, dengan pertolongan Allah SWT yang datang lewat dukungan dari berbagai pihak, baik secara langsung maupun tidak langsung sehingga semuanya dapat teratasi.

Ungkapan rasa terima kasih yang sebesar-besarnya penulis sampaikan kepada orangtua dan suami yang sangat penulis sayangi dan hormati, yaitu Ayahanda Sudirman dan Ibunda Erna serta Kakanda Reski yang tiada henti-hentinya memberikan cinta, nasihat, kasih sayang, semangat, pengorbanan dan dukungan yang diberikan kepada penulis selama menempuh pendidikan, serta doa yang tulus diberikan kepada penulis. Demikian juga untuk saudara-saudaraku Nasri dan Nasru atas semangat dan dukungan yang diberikan kepada penulis selama menempuh pendidikan.

Ucapan terima kasih dan penghargaan yang setinggi-tingginya penulis sampaikan kepada:

1. Bapak Prof. Dr. H. Husain Syam, M.TP., selaku Rektor Universitas Negeri Makassar.
2. Bapak Prof. Dr. Abdul Rahman, M.Pd., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam UNM.
3. Bapak Dr. Awi, M.Si. selaku ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam UNM.
4. Ibu Wahidah Sanusi, S.Si., M.Si., Ph.D. selaku Ketua Program Studi Matematika Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam UNM.
5. Bapak Dr. Hisyam Ihsan, M.Si. selaku pembimbing I dan Ibu Hj. Wahidah Sanusi, S.Si., M.Si., Ph.D. selaku pembimbing II dan atas segala bimbingan, masukan, motivasi dan nasehat yang diberikan kepada penulis dalam menyelesaikan skripsi ini.
6. Bapak Syafruddin Side, S.Si., M.Si., Ph.D. selaku Penguji I dan Bapak Ahmad Zaki, S.Si., M.Si. selaku Penguji II yang telah banyak memberikan saran kepada penulis.

7. Bapak dan Ibu Dosen Jurusan Matematika FMIPA UNM atas segala bantuan dan bekal ilmu pengetahuan yang diberikan kepada penulis.
8. Kak Adiatma yang telah memberi semangat dan bimbingan kepada penulis selama menyelesaikan skripsi ini.
9. Keluarga besar ayah dan ibu serta para sepupu yang senantiasa mendoakan dan memberi semangat kepada penulis, serta memberikan banyak bantuan kepada penulis selama menjalani masa studi di Jurusan Matematika FMIPA UNM.
10. Teman-teman terbaik, tercinta dan sekaligus seperjuangan "Matematika Sains 2013" Hikma, Erna, Ayu, Wakiah, Ida, Meisy, Noni, Gita, Diah, Eni, Amma, Sukma, Diki, Selvi, Yanti, Anto, gusman, Ilham, Ica, Sella, Rahmah, Edi, Raid, Aswar, Titi, Wati, Pute, Disa, Dillah, Dayat, Mimin, Eka, Arif, Katrin, Anti, Imam, Wawan, Qadri, Odi, Iski, April, Taslim, dan Rahmat.
11. Terkhusus penulis ucapkan terima kasih kepada Nur Hikmayanti Syam, Andi Ayu Nurhidayah, dan Herna, terima kasih telah menjadi teman dekat dan terbaik penulis selama menjalankan Studi di Jurusan Matematika FMIPA UNM.
12. Kepada teman-teman "KKN PPM Kab. Pinrang 2016 terkhusus desa Lanrisang" Ida, Gita, Eni, Raid, Pute, Dayat, Arif, Qadri, April, Hilma, Dian, Kak Alim, Kak Trys, Haidir, Ikky, Ana, Tri, Qalbi, Ute, Fitri, Tami, Bella, Taswin, Joni, dan Erwin. Terima kasih telah mengukir kenangan yang baik kepada penulis selama menjalani masa KKN PPM
13. Terima kasih kepada kakak-kakak alumni Matematika Sains dan juga kepada adik-adik junior Matematika FMIPA UNM.

Banyak insan yang telah berjasa dalam hidup ini, tapi lembaran-lembaran ini tidaklah cukup untuk semuanya. Akhirnya penulis berharap semoga bantuan yang telah diberikan mendapatkan balasan dari Allah SWT sebagai amal jariyah dan pahala yang berlipat ganda di sisi-Nya. Aamiin Ya Rabbal Alaamiin.

Makassar, Juni 2017

Penulis

DAFTAR ISI

DAFTAR TABEL

DAFTAR SIMBOL

Y	: variabel respon
X	: variabel prediktor
$Var\ (Y)$	: variansi (Y)
$E(Y)$	: nilai tengah (Y)
k	: parameter dispersi
μ_i	: laju kejadian
e	: nilai euler
β	: parameter koefisien
η	: penduga linear
$g(.)$	: fungsi penghubung
$L(\widehat{\omega})$	: nilai maximum likelihood model sederhana
$L(\widehat{\Omega})$	: nilai maximum likelihood model lengkap
D	: nilai uji Devians
W	: nilai uji Wald
α	: taraf signifikansi
VIF	: *variance inflation factors*
ε	: galat error
$SE(\beta)$	: standar error dari β
db	: derajat bebas

DAFTAR LAMPIRAN

BAB I

PENDAHULUAN

A. Latar Belakang

Ilmu pengetahuan secara umum adalah suatu sistem berbagai pengetahuan yang didapatkan dari hasil pemeriksaan-pemeriksaan yang dilakukan secara teliti dengan menggunakan metode-metode tertentu (Syafiee, 2005). Ilmu pengetahuan dikembangkan sebagai suatu usaha untuk menjelaskan berbagai fenomena yang ada di alam. Matematika merupakan ilmu yang mendasari dan berkaitan dengan ilmu pengetahuan lain. Matematika merupakan ilmu yang sangat penting karena sangat dibutuhkan dalam kehidupan sehari-hari.

Analisis regresi merupakan salah satu cabang dari ilmu matematika yang sering digunakan dalam menyelesaikan suatu masalah. Analisis regresi adalah analisis yang digunakan untuk menjelaskan keterkaitan hubungan antara suatu variabel bebas (variabel prediktor) terhadap variabel tak bebas (variabel respon) (Krisnawardhani, 2010). Pada umumnya analisis regresi digunakan untuk menganalisis data variabel respon yang berupa data kontinu. Namun, dalam pengaplikasiannya data variabel respon yang akan dianalisis juga dapat berupa data diskrit (data *count*) (Ruliana, 2015). Salah satu bentuk paling umum dari *Count Regression* adalah regresi Poisson.

Regresi poisson merupakan metode regresi yang digunakan untuk menganalisis data yang variabel responnya berupa data diskrit (Putra, 2013). Regresi Poisson digunakan untuk menganalisis hubungan antara variabel respon (Y) yang berupa data diskrit dengan beberapa variabel prediktor (X) yang dapat berupa data diskrit, kontinu, kategorik, ataupun campuran.

Terdapat asumsi yang harus dipenuhi dalam regresi Poisson, yaitu nilai varians dan nilai rata-rata dari variabel respon Y harus memiliki nilai yang sama atau *equidispersi*. Namun, dalam kenyataan di lapangan sering terjadi pelanggran asumsi tersebut, yaitu nilai variansnya lebih besar dari nilai rata-rata yang dinamakan *overdispersi* atau nilai variansnya lebih kecil dari nilai rata-rata yang dinamakan *underdispersi* (Wang & Famoye, dalam Putra, 2013)). Menggunakan analisis regresi Poisson saat terjadi *overdispersi* pada data menjadi kurang akurat karena berdampak pada nilai *standard error* menjadi *underestimate* (lebih kecil dari nilai sesungguhnya) sehingga menghasilkan kesimpulan yang tidak valid. Untuk menangani masalah *overdispersi* atau *underdispersi* pada regresi Poisson, terdapat berbagai pilihan model regresi yang dapat digunakan, salah satu diantaranya adalah analisis *Generalized Poisson Regression* yang merupakan perluasan dari regresi Poisson. (Putra, 2013).

Aulele (2012) telah melakukan pemodelan menggunakan metode regresi Poisson pada kasus jumlah kematian bayi di Provinsi Maluku. Penelitian lain yang juga menerapkan regresi Poisson sebagai penyelesaian masalah dalam penelitiannya dilakukan oleh Yulianingsih (2012) untuk mengetahui faktor-faktor yang mempengaruhi jumlah siswa tidak lulus ujian nasional di Provinsi Bali untuk tingkat SMA/SMK. Namun, tidak semua data yang diteliti menggunakan regresi Poisson memenuhi asumsi *equidispersi* seperti yang diharuskan pada regresi Poisson. Sehingga, Cahyandari (2014) melakukan pengujian *overdispersi* pada regresi Poisson. Penelitian lain juga dilakukan oleh Putra (2013) dengan menerapkan *Generalized Poisson Regression* untuk mengatasi fenomena *overdispersi* pada regresi Poisson.

Model *Generalized Poisson Regression* merupakan perluasan dari model regresi Poisson yang digunakan untuk data diskrit ketika terjadi pelanggaran asumsi *equidispersi* pada regresi Poisson. Hubungan nilai varians dan nilai rata-rata dalam *Generalized Poisson Regression* dapat dikondisikan sebagai berikut (Putra, 2013):

1) Jika nilai varians sama dengan nilai rata-rata $VarV(Y_i|x_i) = E(Y_i|x_i)$, maka nilai parameter dispersi $k = 0$, sehingga fungsi densitas peluang *Generalized Poisson*, akan diturunkan ke regresi Poisson
2) Jika nilai varians lebih besar dari nilai rata-rata $VarV(Y_i|x_i) > E(Y_i|x_i)$, maka nilai parameter dispersi $k > 0$, sehingga dapat dikatakan pada data terjadi *overdispersi*
3) Jika nilai varians lebih kecil dari nilai rata-rata $VarV(Y_i|x_i) < E(Y_i|x_i)$, maka nilai parameter dispersi $k < 0$, sehingga dapat dikatakan pada data terjadi *underdispersi*.

Jumlah angka putus sekolah merupakan data diskrit yang mengikuti distribusi Poison sehingga untuk mengetahui faktor-faktor yang berpotensi dalam peningkatan jumlah angka putus sekolah dapat dilakukan pemodelan dengan menggunakan analisis regresi Poisson.

Angka putus sekolah mencerminkan persentase anak-anak usia sekolah yang sudah tidak lagi bersekolah atau tidak menamatkan jenjang pendidikan tertentu (Riyadi, dkk., 2015) dan angka partisispasi sekolah merupakan proporsi penduduk kelompok usia tertentu yang masih duduk di bangku sekolah (BPS, 2015).

Hasil perhitungan Badan Pusat Statistik (BPS) menunjukkan rata-rata penduduk Sulawesi Selatan mengenyam pendidikan hanya sampai kelas 1-2 SLTP. Pemerintah Provinsi Sulawesi Selatan terus berupaya meningkatkan pendidikan masyarakatnya dengan program wajib belajar 12 tahun dan pendidikan gratis. Upaya ini memberikan hasil yang baik dimana harapan lama sekolah setiap orang mencapai 12,9 tahun dan penduduk usia 7-15 tahun hampir seluruhnya bersekolah. Capaian pendidikan di Sulawesi Selatan semakin meningkat selama kurun waktu 2010. Namun demikian, grafik yang menunjukkan partisipasi sekolah menurut umur di Sulawesi Selatan mulai melandai pada umur 13 tahun yang berarti masih ada anak yang tidak melanjutkan sekolah karena berbagai hal (BPS, 2015).

Hal serupa juga dapat dilihat dari sisi pengeluaran konsumsi masyarakat Sulawesi Selatan, mayoritas penduduk usia 7-12 tahun dari kalangan masyarakat miskin sampai kaya, sedang bersekolah, sedang pada penduduk usia selanjutnya 13-15 tahun dan 16-18 tahun, penduduk dengan pengeluaran terendah semakin menunjukkan banyak dari mereka yang mengalami putus sekolah (BPS, 2015).

Putus sekolah merupakan salah satu indikator yang berguna untuk mengukur kemajuan sumber daya manusia pada bidang pendidikan. Oleh karena itu, masalah anak putus sekolah tersebut perlu mendapatkan perhatian. Untuk menekan laju pertumbuhan jumlah anak putus sekolah dapat dilakukan dengan cara mengetahui faktor-faktor yang berpotensi memberikan sumbangsih terhadap pertumbuhan jumlah anak putus sekolah. Oleh kerena itu, peneliti tertarik untuk mengkaji faktor-faktor yang mempengaruhi banyaknya kasus putus sekolah usia wajib belajar di Provinsi Sulawesi Selatan tahun 2016. Sehingga variabel respon yang digunakan yaitu jumlah angka putus sekolah pada usia wajib belajar di Provinsi Sulawesi Selatan tahun 2015 dan variabel prediktor yang digunakan yaitu rasio siswa terhadap sekolah, rasio siswa terhadap guru, rata-rata anggota rumah tangga, jumlah penduduk miskin, kepadatan penduduk, dan rata-rata lama bersekolah di Provinsi Sulawesi Selatan.

Berdasarkan latar belakang tersebut, sehingga judul yang akan dikaji dalam skripsi ini adalah “Pemodelan Angka Putus Sekolah bagi Anak Usia Wajib Belajar Di Provinsi Sulawesi Selatan dengan Pendekatan *Generalized Poisson Regression* (GPR)”.

B. Rumusan Masalah

Berdasarkan latar belakang tersebut maka rumusan masalahnya, adalah :

1. Bagaimana menaksir parameter model *Generalized Poisson Regression* menggunakan metode *Maximum Likelihood Estimation*?

2. Bagaimana model angka putus sekolah di Sulawesi Selatan menggunakan analisis model *Generalized Poisson Regression*?
3. Faktor-faktor apa saja yang berpengaruh secara signifikan terhadap pertumbuhan angka putus sekolah di Provinsi Sulawesi Selatan?

C. Tujuan

Dengan memperhatikan rumusan masalah maka tujuan yang ingin dicapai adalah :

1. Mengetahui cara untuk menafsirkan parameter model *Generalized Poisson Regression* menggunakan metode *Maximum Likelihood Estimation.*
2. Mendapatkan model *Generalized Poisson Regression* pada data banyaknya angka putus sekolah yang terjadi di Provinsi Sulawesi Selatan.
3. Mengetahui faktor-faktor yang memberikan pengaruh signifikan terhadap pertumbuhan angka putus sekolah yang terjadi di Provinsi Sulawesi Selatan.

D. Manfaat

1. Manfaat Teoritis

   Hasil penelitian ini dapat dijadikan bahan studi lanjutan yang relevan dan sebagai bahan kajian kearah pengembangan. Penelitian ini juga dapat dijadikan bahan bacaan untuk menambah wawasan mengenai penerapan matematika dibidang sosial maupun dibidang ilmu lainnya.

2. Manfaat Praktis

   Penelitian ini secara praktis diharapkan dapat memiliki konstribusi sebagai berikut :

a. Bagi Pemerintah

Hasil penelitian ini dapat dijadikan sebagai dasar perencanaan untuk menangani masalah putus sekolah dengan lebih fokus terhadap faktor-faktor yang memberikan pengaruh terhadap peningkatan angka putus sekolah.

b. Bagi penulis

Hasil penelitian ini dapat dijadikan bahan untuk melakukan penelitian lebih lanjut yang relevan.

c. Bagi Pembaca

Hasil penelitian ini diharapkan dapat dijadikan sebagai bahan bacaan dan menjadi perbandingan bagi pembaca yang sedang melakukan penelitian.

BAB II

KAJIAN PUSTAKA

A. Distribusi Pisson

1. Pengertian Distribusi Poisson

Percobaan yang menghasilkan peubah acak Y yang bernilai numerik, yaitu banyaknya hasil selama selang waktu tertentu atau dalam daerah tertentu disebut percobaan Poisson. Banyaknya nilai Y dalam suatu percobaan Poisson disebut suatu peubah acak Poisson dan distribusi peluangnya disebut distribusi Poisson. Distribusi Poisson merupakan suatu distribusi dimana kejadian tergantung pada interval waktu tertentu atau disuatu daerah tertentu dengan hasil pengamatan berupa variabel diskrit (Walpole, 1995).

Definisi 2.1: ***Distribusi Poisson*** *Distribusi peluang peubah acak Y yang menyatakan banyaknya sukses yang terjadi dalam suatu selang waktu atau daerah tertentu dinyatakan dengan i diberikan oleh persamaan* (2.1)

$$f(y;\mu_i) = \frac{e^{-\mu_i}{\mu_i}^y}{y!} \qquad y_i = 0,1,2,\ldots \tag{2.1}$$

Keterangan:

$\mu_i =$ Laju kejadian (rata-rata banyaknya sukses yang terjadi per satuan waktu atau daerah tersebut)

$y =$ Jumlah pengamatan yang dilakukan pada waktu ke-i

$e = 2{,}71878$......

B. Regresi Poisson

Regresi Poisson merupakan model regresi non-linear yang sering digunakan untuk menganalisis suatu data diskrit. Regresi Poisson mengacu pada penggunaan distribusi Poisson (Myers, 1990). Pada regresi Poisson terdapat asumsi, yaitu variabel respon (Y) berdistribusi Poisson dan tidak terjadi multikolinieritas di antara diantara variabel prediktor (X) (Hinde dalam Astuti, 2013).

Regresi Poisson merupakan penerapan dari *Generalized Linear Model* (GLM) yang menggambarkan hubungan antara variabel respon Y data diskrit berdistribusi Poisson dengan variabel prediktor X yang berupa data diskrit atau kontinu (Simarmata, 2011). *Generalized Linear Model* (GLM) merupakan perluasan dari model regresi umum untuk variabel respon yang memiliki sebaran keluarga eksponensial (Astuti, 2013). Distribusi Poisson juga termasuk dalam keluarga eksponensial, hal ini diperlihatkan oleh fungsi peluang densitasnya sebagaiamana persamaan (2.2) (Agresti, 2002)

$$f(y_i;\mu_i) = \frac{e^{-\mu_i}\mu_i{}^{y_i}}{y_i!} = \frac{\exp(-\mu_i)\exp(y_i \ln \mu_i)}{y_i!} \tag{2.2}$$

Menurut Agresti (2002) terdapat 3 komponen utama dalam *Generalized Linear Model* (GLM), yaitu komponen random, komponen sistematik, dan *link function*. Komponen random diidentifikasikan oleh variabel respon (Y) dengan $E(Y) = \mu$. Komponen random dari *Generalized Linear* Model (GLM) termasuk kedalam keluarga ekponensial yang fungsi peluang densitasnya dinyatakan pada persamaan (2.3)

$$f(y_i,\theta) = a(\theta)b(y_i)\exp\big(y_i Q(\theta)\big) \tag{2.3}$$

Berdasarkan persamaan (2.2) dan persamaan (2.3) maka dapat dinyatakan bahwa, $\theta = \mu_i, a(\mu) = \exp(\mu_i), b(y) = \frac{1}{y_i!}$, dan $Q(\mu) = \ln(\mu_i)$, sehingga terbukti bahwa Poisson termasuk dalam keluarga eksponensial.

Komponen sistematik meliputi variabel-variabel penjelas dari model. Komponen sistematik merupakan kombinasi linear variabel prediktor $\boldsymbol{X}$ dengan koefisien parameter $\boldsymbol{\beta}$ yang menghasilkan penduga linear $\boldsymbol{\eta}$, yaitu $\boldsymbol{\eta} = \boldsymbol{X\beta}$. Masing-masing dari elemen $\boldsymbol{\eta}$ dapat dinyatakan dengan $\eta_i = \sum_{j=0}^{k} \beta_j x_{ij}$.

Komponen ketiga, yaitu *link function* $g(.)$. *Link Function* menghubungkan komponen random dan komponen sistematik. *Link function* merupakan fungsi yang menghubungkan mean variabel respon Y (μ) dengan variabel-variabel penjelas atau penduga linear (η) dan diformulasikan sebagaimana persamaan (2.4)

$$g(\mu_i) = \eta_i = \sum_{j=0}^{k} \beta_j x_{ij} = \beta_0 + \beta_i x_{i1} + \cdots + \beta_k x_{ik} \tag{2.4}$$

fungsi g tersebut merupakan fungsi penghubung. Sedangkan hubungan antara mean dan penduga linear adalah seperti pada persamaan (2.5) (Cahyandari, 2014).

$$\mu_i = g^{-1}(\eta_i) = g^{-1}(x_i^T \beta) \tag{2.5}$$

Link function yang biasa digunakan dalam regresi Poisson adalah fungsi penghubung log karena rata-rata dari variabel responnya akan berbentuk fungsi eksponensial dan menjamin bahwa nilai variabel yang ditaksir dari variabel responnya akan bernilai non-negatif. Fungsi penghubung log berbentuk sebagaimana persamaan (2.6) (Cahyandari, 2014)

$$g(\mu_i) = \ln \mu_i = x_i^T \beta \tag{2.6}$$

hubungan antara mean variabel respon dan penduga linear dengan menggunakan fungsi penghubung log diberikan oleh persamaan (2.7)

$$\ln \mu_i = x_i^T \beta$$

$$e^{\ln \mu_i} = e^{x_i^T \beta}$$

$$\mu_i = e^{x_i^T \beta} \tag{2.7}$$

karena *Link function* yang digunakan dalam regresi Poisson adalah ln, maka $ln(\mu_i) = \eta_i$, sehingga fungsi penghubung regresi poisson dapat dinyatakan sebagaimana persamaan (2.8)

$$ln(\mu_i) = x_i^T \beta$$

$$= \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} \qquad , i = 1,2, \cdots, n \ ; j = 1,2, \cdots, p$$

$$= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}$$

$$\mu_i = \exp(x_i^T \beta)$$

$$= \exp\left(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}\right)$$

$$= \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}\right) \tag{2.8}$$

Dalam regresi Poisson terdapat asumsi yang harus dipenuhi, yaitu Variabel respon (Y) merupakan data diskrit serta nilai varians dan nilai rata-rata dari variabel respon (Y) harus memiliki nilai yang sama atau asumsi *equidispersi* sebagaimana persamaan (2.10)

$$Var\ (Y|x) = E(Y|x) = \mu \tag{2.10}$$

jadi, asumsi yang harus terpenuhi pada model regresi Poisson ditunjukkan pada persamaan (2.11) (Fadhillah, 2011)

$$Var(Y_i) = E(Y_i) = \mu_i = \exp(\boldsymbol{x_i^T \beta}) \tag{2.11}$$

dengan $\boldsymbol{x}_i$ adalah vektor yang berukuran $p \times 1$ yang menjelaskan variabel prediktor dan $\boldsymbol{\beta}$ adalah vektor berukuran $p \times 1$ yang merupakan parameter regresi. Oleh karena *link function* pada regresi Poisson menggunakan fungsi penghubung ln maka fungsi kepadatan peluang regresi Poisson adalah sebagaimana persamaan (2.12) (Cahyandari, 2014).

$$f(y_i, \boldsymbol{\beta}) = \frac{{\mu_i}^{y_i} e^{-\mu_i}}{y_i!}$$

$$= \frac{\left[e^{\boldsymbol{x_i^T \beta}}\right]^{y_i} e^{-\left[e^{\boldsymbol{x_i^T \beta}}\right]}}{y_i!} \quad (2.12)$$

Mean dan varians untuk model regresi Poisson adalah sebagai berikut:

Mean = $\mu_i = \exp \boldsymbol{x_i^T \beta} = \exp\left(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}\right)$

Varians = $Var(Y_i) = \exp \boldsymbol{x_i^T \beta} = \exp\left(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}\right)$

Model regresi Poisson dengan penghubung log dapat ditulis sebgaimana persamaan (2.13) (Cahyandari, 2014)

$$y_i = \mu_i + \varepsilon_i = \exp \boldsymbol{x_i^T \beta} + \varepsilon_i \quad (2.13)$$

dengan

$\boldsymbol{x_i} = [1 \quad x_{i1} \quad x_{i2} \quad \cdots \quad x_{ip}]^T \quad \boldsymbol{\beta} = [\beta_0 \quad \beta_1 \quad \beta_2 \quad \cdots \quad \beta_p] \;\; i = 1,2,\cdots,n$

y_i = jumlah kejadian ke-ij

μ_i = mean jumlah kejadian

ε_i = galat error atau residual

1. Penaksiran Parameter pada Model Regresi Poisson

Pada model regresi Poisson harus dilakukan penaksiran parameter yang tidak diketahui $(\boldsymbol{\beta_j})$, $j = 1,2,3,\cdots,p$. Salah satu metode penaksir parameter yang dapat digunakan pada regresi Poisson untuk menaksir parameter suatu model yang diketahui distribusinya adalah metode *Maximum Likelihood Estimation* (MLE) (Cahyandari, 2014).

Berdasarkan persamaan (2.11), maka penaksiran parameter $(\boldsymbol{\beta_j})$ menggunakan metode *Maximum Likelihood Estimation* (MLE) dapat dilakukan dengan langkah-langkah sebagai berikut:

1. Mengambil n buah sampel acak
2. Membentuk fungsi *Likelihood* berdasarkan persamaan (2.12)

   Fungsi *Likelihood* untuk model regresi Poisson diberikan oleh persamaan (2.14) (Putra, 2013)

$$
\begin{aligned}
L(y_i, \boldsymbol{\beta}) &= \prod_{i=1}^{n} f(\, y_i, \boldsymbol{\beta}) \\
&= \prod_{i=1}^{n} \frac{\mu_i{}^{y_i} e^{-\mu_i}}{y_i!} \\
&= \frac{\{\prod_{i-1}^{n} \mu_i{}^{y_i}\} \prod_{i=1}^{n} e^{-\mu_i}}{\prod_{i=1}^{n} y_i!} \\
&= \frac{\{\prod_{i-1}^{n} \mu_i{}^{y_i}\} e^{-\sum_{i=1}^{n} u_i}}{\prod_{i=1}^{n} y_i!} \qquad (2.14)
\end{aligned}
$$

3. Mengambil bentuk log dari fungsi *Likelihood* yang diperoleh.

   Persamaan (2.14) dibentuk menjadi fungsi log-*likelihood* agar mudah diselesaikan. Fungsi log-*likelihood* yang terbentuk ditunjukkan oleh persamaan (2.15) (Ruliana, 2015)

$$
\begin{aligned}
\ln L(y_i, \boldsymbol{\beta}) &= \ln \left\{ \frac{\{\prod_{i-1}^{n} \mu_i{}^{y_i}\} e^{-\sum_{i=1}^{n} u_i}}{\prod_{i=1}^{n} y_i!} \right\} \\
&= \ln \left\{ \prod_{i=1}^{n} \left(\frac{\mu_i{}^{y_i} e^{-\mu_i}}{y_i!} \right) \right\} \\
&= \ln \left\{ \prod_{i=1}^{n} \left\{ \frac{\left[e^{\boldsymbol{x_i^T \beta}}\right]^{y_i} e^{-\left[e^{\boldsymbol{x_i^T \beta}}\right]}}{y_i!} \right\} \right\}
\end{aligned}
$$

$$= \ln\left\{\prod_{i=1}^{n}\left\{\frac{\left[e^{\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}}\right]^{y_i} e^{-\left[e^{\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}}\right]}}{y_i!}\right\}\right\}$$

$$= \sum_{i=1}^{n}\ln\left\{\frac{\left[e^{\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}}\right]^{y_i} e^{-\left[e^{\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}}\right]}}{y_i!}\right\}$$

$$= \sum_{i=1}^{n}\ln\left\{\frac{\left[e^{\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}\right)^{y_i}}\right] e^{-\left[e^{\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}}\right]}}{y_i!}\right\}$$

$$= \sum_{i=1}^{n}\left\{\ln\left[e^{\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}\right)^{y_i}}\right] + \ln\left(e^{-\left[e^{\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}}\right]}\right) - \ln(y_i!)\right\}$$

$$= \sum_{i=1}^{n}\left\{y_i \ln\left(e^{\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}\right)}\right) - \left[e^{\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}}\right] - \ln(y_i!)\right\}$$

$$= \sum_{i=1}^{n}\left\{y_i\left(\beta_0 + \sum_{j=1}^{p}\beta_j x_{ij}\right) - e^{\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}\right)} - \ln(y_i!)\right\} \quad (2.15)$$

4. Mendiferensialkan fungsi log-*likelihood* terhadap masing-masing parameter $\boldsymbol{\beta}$

   Persamaan (2.15) didiferensialkan terhadap masing-masing parameter $\left(\beta_j\ , j = 0,1,2,\cdots,p\right)$ dan disamakan dengan nol sehingga akan diperoleh taksiran parameter yang maksimum atas penyelesaian persamaan (2.15) sebagaimana pada persamaan (2.16) (Pratiwi, 2012)

$$R(\boldsymbol{\beta}) = \begin{bmatrix} R_0(\boldsymbol{\beta}) \\ R_1(\boldsymbol{\beta}) \\ \vdots \\ R_p(\boldsymbol{\beta}) \end{bmatrix} = \begin{bmatrix} \frac{\partial \ln L(y_i,\boldsymbol{\beta})}{\partial \beta_0} \\ \frac{\partial \ln L(y_i,\boldsymbol{\beta})}{\partial \beta_1} \\ \vdots \\ \frac{\partial \ln L(y_i,\boldsymbol{\beta})}{\partial \beta_p} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^{n}\{y_i - \exp(\beta_0 + \sum_{j=1}^{p}\beta_j x_{ij})\} \\ \sum_{i=1}^{n}\{x_{i1}(y_i - \exp(\beta_0 + \sum_{j=1}^{p}\beta_j x_{ij}))\} \\ \vdots \\ \sum_{i=1}^{n}\{x_{ip}(y_i - \exp(\beta_0 + \sum_{j=1}^{p}\beta_j x_{ij}))\} \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$= 0 \tag{2.16}$$

penentuan penaksir titik $\boldsymbol{\beta}$ secara manual akan mengalami kesulitan, sehingga dalam penentuannya dilakukan dengan menggunakan bantuan *software* statistik. Dalam hal ini, penulis menggunakan *software* SPSS untuk membantu proses perhitungan taksiran nilai $\boldsymbol{\beta}$.

2. Pengujian Parameter Model Regresi Poisson

Sebelum menentukan nilai statistik uji, maka terlebih dahulu dilakukan pengujian parameter pada model regresi Poisson untuk mengetahui ada atau tidak ada pengaruh variabel prediktor terhadap variabel Respon. Pengujian parameter pada model regresi Poisson dilakukan menggunakan Uji Serentak dan Uji Parsial.

1. Uji Serentak Parameter Model Regresi Poisson

Uji serentak parameter model regresi Poisson digunakan untuk mengetahui ada tidaknya pengaruh variabel prediktor terhadap variabel respon. (Darnah dalam Ruliana, 2015). Pengujian parameter model regresi Poisson dilakukan dengan menggunakan metode *Maximum Likelihood Ratio Test* (MLRT). Hal pertama yang dilakukan adalah menentukan dua fungsi likelihood yang berhubungan dengan model regresi yang diperoleh, yaitu $L(\hat{\Omega})$ dan $L(\hat{\omega})$. (Fadmi, 2014).

$L(\hat{\Omega})$ adalah nilai maksimum *likelihood* untuk model lengkap dengan melibatkan variabel prediktor dan $L(\hat{\omega})$ adalah nilai maksimum *likelihood* untuk model sederhana tanpa melibatkan variabel prediktor. (Fadmi, 2014) Setelah memaksimumkan kedua fungsi *likelihood* yang diperoleh maka akan didapatkan nilai *likelihood*-nya.

Hipotesis yang diujikan:

$H_0 : \beta_1 = \beta_2 = \cdots = \beta_p = 0$

$H_1 :$ paling sedikit terdapat satu $\beta_j \neq 0 \,; j = 1, 2, 3, \cdots, p$

taraf signifikansi yang digunakan:

$\alpha = 0{,}05$

statistik uji yang digunakan sebagaimana yang ditunjukkan pada persamaan (2.17) (Ruliana, 2015)

$$G = -2 \ln \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right)$$

$$= 2\left[\ln L(\hat{\Omega}) - \ln L(\hat{\omega})\right] \qquad (2.17)$$

G merupakan deviansi model regresi Poisson atau devians yang dihitung pada seluruh parameter dalam model. Nilai G yang semakin kecil menunjukkan semakin kecil pula tingkat kesalahan yang dihasilkan model sehingga model menjadi semakin tepat (Fadmi, 2014).

Statistik uji rasio *Likelihood* mengikuti sebaran χ^2 (*Chi-Square*) dengan derajat bebasnya adalah banyaknya parameter model dibawah populasi dikurangi dengan banyaknya parameter dibawah H_0 (Fadmi, 2014). Kriteria pengujiannya adalah tolah H_0 jika $G > \chi^2_{(k;\alpha)}$.

2. Uji Parsial Parameter Model Regresi Poisson

Selanjutnya dilakukan pengujian parameter model secara parsial. Pengujian secara parsial digunakan untuk mengetahui parameter mana yang memberikan pengaruh signifikan terhadap model. (Fadmi, 2014). Statistik uji yang digunakan untuk uji parsial, yaitu uji *Wald* dengan hipotesis (Ruliana, 2015):

$H_0 : \beta_j = 0$; untuk suatu $j = 1, 2, \cdots, p$

$H_1 : \beta_j \neq 0$; untuk suatu $j = 1, 2, \cdots, p$

taraf signifikansi yang digunakan, yaitu $\alpha = 0{,}05$

statistik uji Wald ditunjukkan oleh persamaan (2.18)

$$W = \left(\frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}\right)^2 \tag{2.18}$$

dengan

$\hat{\beta}_j$ = taksiran parameter β_j

$SE(\hat{\beta}_j)$ = adalah taksiran standar error dari β_j

kriteria pengujian:

Tolak H_0 jika atau tolak H_0 jika $W > \chi^2{}_{(\alpha,1)}$, α merupakan taraf signifikansi yang digunakan.

C. Uji Ketepatan (*Goodness of Fit*) Model Regresi Poisson

Uji *Goodness of Fit* bertujuan untuk mengetahui model yang digunakan sesuai atau tidak dengan data yang diamati (Cahyandari, 2014). Uji ini didasarkan pada seberapa baik kesesuain/kecocokan antara frekuensi pengamatan yang diperoleh data sampel dengan frekuensi harapan yang diperoleh dari distribusi yang dihipotesiskan (Lungan dalam Ruliana, 2015).

Uji ketepatan pada model regresi Poisson dapat dilakukan dengan dua cara, yaitu *Pearson Chi-Square Statistic* dan *Deviance* (Irwan, 2013). Dalam hal ini penulis memilih untuk menggunakan Devians dalam uji ketepatan model regresi Posisson. Devians dapat diartikan sebagai logaritma dari uji *Likelihoodnya*, yaitu seperti pada persamaan (2.19) (Irwan, 2013).

$$D = -2\ln\left[\frac{L(y,\mu)}{L(y,y)}\right]$$

$$= 2[\ln L(y,y) - \ln L(y,\mu)] \tag{2.19}$$

Pada persamaan (2.19), $L(y,\mu)$ adalah fungsi *likelihod current* model sedangkan $L(y,y)$ adalah fungsi *likelihood saturated* model dari distribusi Poisson. Adapun fungsi log *likelihood current* modelnya ditunjukkan oleh persamaan (2.20) (Irwan, 2013).

$$L(y,\mu) = \ln L(y,\mu)$$

$$= \ln \prod_{i=1}^{n} \frac{e^{-\mu_i}\mu_i{}^{y_i}}{y_i!}$$

$$= \sum_{i=1}^{n} (y_i \ln \mu_i - \mu_i - \ln y_i!) \quad (2.20)$$

Sedangkan untuk fungsi log *likelihood saturated* model, nilai-nilai μ_i diganti dengan nilai y_i (tanpa asumsi tentang keeratan hubungannya dengan variabel x nya). Fungsi log *likelihood saturated* modelnya ditunjukkan pada persamaan (2.21)

$$L(y, y) = \ln L(y, y)$$

$$= \ln \prod_{i=1}^{n} f(y_i, y_i)$$

$$= \ln \prod_{i=1}^{n} \frac{e^{-y_i} {y_i}^{y_i}}{y_i!}$$

$$= \sum_{i=1}^{n} (y_i \ln y_i - y_i - \ln y_i!) \quad (2.21)$$

Dengan demikian nilai Devians bisa diperoleh dengan mensubstitusi persamaan (2.20) dan (2.21) ke dalam persamaan (2.19) sehingga akan diperoleh persamaan (2.22) sebagai bentuk dari uji statistik Devians (Irwan, 2013)

$$D = 2(\ln L(y, y) - \ln L(y, \mu))$$

$$= 2\{\sum_{i=1}^{n} (y_i \ln y_i - y_i - \ln y_i!) - \sum_{i=1}^{n} (y_i \ln \mu_i - \mu_i - \ln y_i!)\}$$

$$= 2 \sum_{i=1}^{n} \left\{ y_i \ln \frac{y_i}{\mu_i} - (y_i - \mu_i) \right\}$$

karena $\sum_{i=1}^{n} (y_i - \mu_i) = 0$ maka akan menjadi sebagaimana persamaan (2.22) (Irwan, 2013).

$$D = 2 \sum_{i=1}^{n} \left\{ y_i \ln \frac{y_i}{\mu_i} \right\} \quad (2.22)$$

Nilai *D* pada persamaan (2.22) disebut statistik uji atau Devians untuk model regresi Poisson. Untuk model yang sesuai, devians mendekati distribusi Chi-Kuadrat (Cahyandari, 2014)

hipotesis yang diujikan, yaitu:

$H_0 =$ Sampel berasal dari populasi yang berdistribusi Poisson

$H_1 =$ Sampel tidak berasal dari populasi yang berdistribusi Poisson

taraf signifikansi

$\alpha = 0{,}05$

kriteria pengujiannya adalah, tolak H_0 pada taraf signifikansi α, jika $D > \chi^2{}_{(n-(p+1)),\alpha}$.

D. Parameter Dispersi

Parameter dispersi (k) diperoleh dari rumus pada persamaan (2.23) (Darnah dalam Ruliana, 2015)

$$k = \frac{D}{db} \tag{2.23}$$

$D =$ niali Devians

$db =$ derajat bebas

Apabila nilai $k > 0$ maka terjadi *overdispersi* dan apabila $k < 0$ maka terjadi *underdispersi*.

E. Overdispersi dan Underdispersi

Pada regresi Poisson terdapat asumsi yang harus dipenuhi, yaitu nilai variansnya sama dengan nilai rata-ratanya atau dikenal dengan istilah *equidispersi*. Namun demikian, seringkali terjadi pelanggaran asumsi pada regressi Poisson. Tidak jarang nilai variansnya melebihi nilai rata-rata atau kurang dari nilai rata-rata. Kondisi ini biasa disebut dengan *overdispersi* dan *underdispersi*.

Overdispersi adalah kondisi ketika data variabel respon menunjukkan nilai variansi lebih besar dari nilai rata-ratanya, sedang *underdipersi* adalah kondisi ketika data variabel respon menujukkan nilai variansi lebih kecil dari nilai rata-ratanya (Darnah dalam Ruliana, 2015). Ketika terjadi fenomena *overdispersi* pada data maka regresi Poisson kurang akurat digunakan untuk analisis karena berdampak pada nilai *standard error* menjadi *under estimate* (lebih kecil dari nilai sesungguhnya), sehingga kesimpulan yang diperoleh menjadi tidak valid (Putra, 2013).

Jika model Poisson yang mengindikasikan gejala *overdispersi* tetap digunakan untuk menggambarkan model data maka taksiran parameter regresinya tetap konsisten tetapi tidak efisien. Akibatnya, taksiran parameter menjadi bias sehingga mengakibatkan kesalahan penarikan kesimpulan mengenai data yang diamati (Pratiwi, 2012). Oleh karena itu dibutuhkan sebuah model yang dapat mengatasi gejala *overdispersi* dan *underdispersi*, salah satu model regresi yang dapat digunakan, yaitu *Generalized Poisson Regression*.

F. AIC (Akaike Information Criterion)

Terdapat beberapa metode dalam menetukan model terbaik, salah satu diantaranya adalah *Akaike Information Criterion* (AIC) (Fadmi, 2014). AIC (*Akaike Information Criterion*) atau “Kriteria Informasi” adalah kriteria untuk memilih model dalam ekonometrika (Ruliana, 2015). AIC digunakan untuk memperoleh model terbaik dari beberapa metode dengan asumsi Poisson. AIC didefinisikan sebagaimana persamaan (2.24) (Bozdogan dalam Fadmi, 2014):

$$AIC = -2 \ln L(\hat{\theta}) + 2p \tag{2.24}$$

dengan

$L(\hat{\theta})$ adalah nilai likelihood dan p adalah jumlah parameter

Model terbaik adalah model yang memiliki nilai AIC terkecil

G. Multikolinieritas

Multikolinieritas berarti terdapat hubungan antara veriabel prediktor yang satu dengan variabel prediktor yang lain (Aulele, 2012). Salah satu syarat yang harus dipenuhi dalam model regresi dengan beberapa variabel prediktor adalah tidak terjadi kasus multikolinieritas atau tidak adanya korelasi antara satu variabel prediktor dengan variabel prediktor lainnya. Apabila kasus ini terjadi maka dapat mengakibatkan nlilai *standard error* yang sangat besar atau matriks $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ memiliki determinan sama dengan nol (Fadmi, 2014).

Untuk mendeteksi ada tidaknya multikolinieritas dapat dilakukan dengan melihat nilai *Tolerance* dan VIF (*Variance Inflation Factor*). Jika nilai *Tollerance* lebih dari 0,1 dan VIF kurang dari 10 maka tidak terjadi multikolinieritas (Priyatno dalam Fadmi, 2014)

hipotesis :

H_0 : Model regresi memiliki masalah multikolinieritas

H_1 : Model regresi tidak memiliki masalah multikolinieritas

taraf signifikansi:

$\alpha = 0{,}05$

statistik uji ditunjukkan oleh persamaan (2.25) dan persamaan (2.26) :

$$VIF = \frac{1}{(1-r_{i,j}{}^2)} \tag{2.25}$$

$$Tollerance = \frac{1}{VIF_j} \quad (2.26)$$

dengan

$r_{i,j}$ = koefisien korelasi antara X_i dan X_j

kriteria uji:

Tolak H_0 jika seluruh variabel prediktor memiliki nilai VIF kurang dari 10 dan nilai *Tolerance* lebih dari 0,1. Sebaliknya, jika seluruh variabel prediktor memiliki nilai VIF lebih dari 10 dan nilai *Tolerance* kurang dari 0,1 maka terima H_0.

Terdapat beberapa cara untuk menghilangkan multikolinieritas, diantaranya yaitu mengeluarkan variabel yang mengandung multikolinieritas, membentuk variabel baru menggunakan PCA (*Principal Component Analysis*), menggunakan metode *ridge* dan masih banyak metode-metode yang dapat digunakan untuk menghilangkan multikolinieritas (Rohmah, 2013).

H. Generalized Poisson Regression

Regresi Poisson tidak sesuai dalam memodelkan data yang mengalami *overdispersi* atau *underdispersi*. Oleh karena itu, dilakukan pendekatan dengan model regresi yang lebih sesuai, salah satu model regresi yang dapat digunakan adalah model *Generalized Poisson Regression* (GPR). Model *Generalized Poisson Regression* (GPR) adalah suatu model regresi yang digunakan untuk menganalisis hubungan antara variabel respon yang berupa data diskrit dengan satu atau lebih variabel prediktor (Famoye dalam Fadmi, 2014).

Generalized Poisson Regression merupakan perluasan dari model regresi Poisson (Pratiwi, 2012). Fungsi peluang *Generalized Poisson Regression* diperlihatkan oleh persamaan (2.27) (Putra , 2013).

$$f(y_i,\mu_i,k) = \left(\frac{\mu_i}{1+k\mu_i}\right)^{y_i} \frac{(1+ky_i)^{y_i-1}}{y_i!} \exp\left(-\frac{\mu_i(1+ky_i)}{(1+k\mu_i)}\right) \tag{2.27}$$

dengan rata-rata dan variansnya masing-masing adalah μ_i dan $\mu_i(1+k\mu_i)^2$. Untuk membuktikan rata-rata dan varians dari *Generalized Poisson Regression*, terlebih dahulu akan dibuktikan bahwan *Generalized Poisson Regression* termasuk dalam distribusi keluarga eksponensial (Pratiwi, 2012).

Fungsi peluang *Generalized Poisson Regression* pada persamaan (2.27) dapat ditulis dalam bentuk seperti pada persamaan (2.28)

$$\begin{aligned} f(y_i,\mu_i,k) &= \exp\left\{\left\{y_i \ln\left(\frac{\mu_i}{1+k\mu_i}\right) - \frac{\mu_i(1+ky_i)}{(1+k\mu_i)} + (y_i-1)(y_i-1)\ln(1+ \right.\right. \\ &\quad \left. ky_i) - \ln y_i!\right\} \\ &= \exp\left\{\left\{y_i \ln\left(\frac{\mu_i}{1+k\mu_i}\right) - \frac{\mu_i}{(1+k\mu_i)} - \frac{k\mu_i y_i}{(1+k\mu_i)} + (y_i-1)\ln(1+ \right.\right. \\ &\quad \left. ky_i) - \ln y_i!\right\} \\ &= \exp\left\{\left\{y_i\left(\ln\left(\frac{\mu_i}{1+k\mu_i}\right) - \frac{k\mu_{i_i}}{(1+k\mu_i)}\right) - \frac{\mu_i}{(1+k\mu_i)} + (y_i - \right.\right. \\ &\quad \left. 1)\ln(1+ky_i) - \ln y_i!\right\} \end{aligned}$$

(2.28)

persamaan (2.28) menunjukkan bahwa *Generalized Poisson Regression* termasuk dalam distribusi keluarga eksponensial dengan:

$$\theta = \ln\left(\frac{\mu_i}{1+k\mu_i}\right) - \frac{k\mu_{i_i}}{(1+k\mu_i)}$$

$$b(\theta) = \frac{\mu_i}{(1+k\mu_i)}$$

$$a(\phi) = 1$$

$$h(y,\phi) = (y_i - 1)\ln(1 + ky_i) - \ln y_i!$$

Selanjutnya akan dicari rata-rata dan variansi *Generalized Poisson Regression*. Rata-rata dan variansi dari keluarga eksponensial adalah $b'\theta$ dan $b''(\theta)a(\phi)$, sehingga dimisalkan rata-rata dan variansi untuk *Generalized Poisson Regression* masing-masing adalah $E(y) = b'\theta$ dan $Var(y) = b''(\theta)a(\phi)$.

Diketahui:

$$\theta = \ln\left(\frac{\mu_i}{1+k\mu_i}\right) - \frac{k\mu_{i_i}}{(1+k\mu_i)}$$

dan

$$b(\theta) = \frac{\mu_i}{(1+k\mu_i)}$$

sehingga

$$\theta = \ln\big(b(\theta)\big) - k\, b(\theta)$$

$$\frac{\partial\theta}{\partial\theta} = \frac{1}{b\theta}\ b'(\theta) - kb'(\theta)$$

$$1 = \frac{b'(\theta) - kb'(\theta)b\theta}{b\theta}$$

$$b\theta = b'(\theta) - kb'(\theta)b\theta$$

$$b\theta = b'(\theta)[1 - kb\theta]$$

$$b'\theta = \frac{b(\theta)}{1-kb\theta}$$

$$= \frac{\frac{\mu_i}{(1+k\mu_i)}}{1-k\frac{\mu_i}{(1+k\mu_i)}}$$

$$= \frac{\frac{\mu_i}{(1+k\mu_i)}}{\frac{(1+k\mu_i)-k\mu_i}{(1+k\mu_i)}}$$

$$= \mu_i \tag{2.29}$$

Berdasarkan persamaan (2.29) diperoleh $E(y) = b'\theta = \mu_i$, dengan demikian terbukti bahwa rata-rata dari *Generalized Poisson Regression* adalah μ_i.

Dari proses pembuktian rata-rata *Generalized Poisson Regression* pada persamaan (2.29) diperoleh persamaan (2.30):

$$1 = \frac{b'(\theta) - kb'(\theta)b\theta}{b\theta}$$

$$= \frac{b'(\theta)}{b\theta} - kb'(\theta)$$

Jika didiferensalkan terhadap θ akan menjadi :

$$0 = \frac{b''(\theta)b(\theta) - \left(b'(\theta)\right)^2}{\left(b(\theta)\right)^2} - kb''(\theta)$$

$$0 = \frac{b''(\theta)b(\theta) - \left(b'(\theta)\right)^2 - kb''(\theta)\left(b(\theta)\right)^2}{\left(b(\theta)\right)^2}$$

$$b''(\theta)b(\theta) - \left(b'(\theta)\right)^2 - kb''(\theta)\left(b(\theta)\right)^2 = 0$$

$$b''(\theta)b(\theta) - kb''(\theta)\left(b(\theta)\right)^2 = \left(b'(\theta)\right)^2$$

$$b''(\theta)\left[b(\theta) - k\left(b(\theta)\right)^2\right] = \left(b'(\theta)\right)^2$$

$$b''(\theta) = \frac{\left(b'(\theta)\right)^2}{b(\theta) - k\left(b(\theta)\right)^2}$$

$$= \frac{(\mu_i)^2}{\frac{\mu_i}{(1+k\mu_i)} - k\left(\frac{\mu_i}{(1+k\mu_i)}\right)^2}$$

$$= \frac{(\mu_i)^2}{\frac{\mu_i}{(1+k\mu_i)} - k\frac{{\mu_i}^2}{(1+k\mu_i)^2}}$$

$$= \frac{(\mu_i)^2}{\frac{\mu_i(1+k\mu_i) - k{\mu_i}^2}{(1+k\mu_i)^2}}$$

$$= \frac{(\mu_i)^2}{\frac{\mu_i + k\mu_i{}^2 - k\mu_i{}^2}{(1+k\mu_i)^2}}$$

$$= \frac{(\mu_i)^2(1+k\mu_i)^2}{\mu_i}$$

$$= \mu_i(1 + k\mu_i)^2 \quad (2.30)$$

Berdasarkan persamaan (2.30) diperoleh $Var(y) = b''(\theta)\alpha(\phi) = \mu_i(1 + k\mu_i)^2 \cdot 1 = \mu_i(1 + k\mu_i)^2$. Dengan demikian, terbukti bahwa variansi dari *Generalized Poisson Regression* adalah $\mu_i(1 + k\mu_i)^2$ (Pratiwi, 2012).

Pada model *Generalized Poisson Regression* terdapat dua paramater, yaitu μ sebagai parameter natural dan k sebagai parameter dispersi (Pratiwi, 2012). Jika nilai parameter dispersi $k = 0$ maka nilai varians sama dengan nilai rata-rata, sehingga fungsi peluang *Generalized Poisson Regression* menjadi fungsi peluang regresi Poisson. Jika nilai parameter dispersi $k > 0$ maka nilai varians lebih besar dari nilai rata-rata sehingga dikatakan terjadi *overdispersi* pada data. Jika nilai parameter dispersi $k < 0$ maka nilai varians lebih kecil dari nilai rata-rata sehingga dikatakan terjadi *underdispersi* pada data (Putra, 2013).

1. Penaksiran Parameter model *Generalized Poisson Regression*

Berdasarkan persamaan (2.7) diketahui bahwa $\mu_i = \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}\right) = \exp\left(\boldsymbol{x_i^T \beta}\right)$, maka fungsi peluang model *Generalized Poisson Regression* dapat dinyatakan sebagaimana persamaan (2.31) (Pratiwi, 2012)

$$f(y_i, \mu_i, k) = \left(\frac{\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}\right)}{1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}\right)}\right)^{y_i} \frac{(1 + ky_i)^{y_i - 1}}{y_i!}$$

$$\exp\left(-\frac{\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}\right)(1 + ky_i)}{\left(1 + k\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}\right)\right)}\right) \quad (2.31)$$

Parameter $\boldsymbol{\beta}$ dan k dapat ditaksir dengan metode *Maximum Likelihood Estimation* (MLE). Pada umumnya penaksiran parameter model *Generalized Poisson Regression* memiliki langkah yang sama dengan penaksiran parameter pada model regresi Poisson, yaitu:

1. Membentuk fungsi Likelihood

   Berdasarkan persamaan (2.31) maka fungsi likelihood untuk model *Generalized Poisson Regression* adalah sebagaimana ditunjukkan pada persamaan (2.32)

   $$L(\boldsymbol{\beta}, k, y) = \prod_{i=1}^{n} f(y_i, \boldsymbol{\beta}, k) \quad (2.32)$$

2. Membentuk fungsi log *likelihood* dari fungsi *likelihood* yang telah diperoleh.

   Fungsi log *likelihood* yang terbentuk dari fungsi *likelihood* pada persamaan (2.32) sebagaimana persamaan (2.33)

   $$\ln L(\boldsymbol{\beta}, k) = \ln \prod_{i=1}^{n} f(y_i, \boldsymbol{\beta}, k) \quad (2.33)$$

3. Mendiferensialkan persamaan log *likelihood* yang telah diperoleh.

   Terdapat tiga langkah dalam mendiferensialkan persamaan log-*likelihood* yang diperoleh pada persamaan (2.31) untuk menaksir parameter model *Generalized Poisson Regression*, yaitu mendiferensialkan persamaan (2.31) terhadap β_0 kemudian disamakan dengan nol, mendiferensialkan persamaan (2.31) terhadap β_j ; $j = 1, 2, \cdots, p$, kemudian disamakan dengan nol, dan mendiferensialkan persamaan (2.31) terhadap k kemudian disamakan dengan nol.

Setelah penaksian parameter selesai, maka diperoleh taksiran model *Generalized Poisson Regression* seperti pada persamaan (2.32) (Putra, 2013).

$$y_i = \exp\left(\boldsymbol{x}_i^T \widehat{\boldsymbol{\beta}}\right) + \varepsilon_i \qquad (2.34)$$

2. Pengujian Parameter model *Generalized Poisson Regression*

Pengujian parameter untuk model *Generalized Poisson regression* dilakukan sama dengan pengujian parameter pada regresi Poisson, yaitu dengan menggunakan metode *Maximum Likelihood Ratio Test* (MLRT) untuk pengujian parameter secara simultan dan menggunakan statistik uji *Wald* untuk pengujian parameter secara parsial (Fadmi, 2014).

a. Pengujian meggunakan statistik uji *Wald* diberikan oleh persamaan (2.35) (Pratiwi, 2012)

hipotesis pengujian :

$H_0 : \beta_j = 0$

$H_1 : \beta_j \neq 0$

statistik uji:

$$W = \left(\frac{\widehat{\beta}_j}{SE(\widehat{\beta}_j)}\right)^2 \qquad (2.35)$$

kriteria pengujian, tolak H_0 jika nilai $W > \chi^2{}_{(\alpha,1)}$ atau tolak H_0 jika nilai $p - value > \alpha$.

b. *Maximum Likelihood Ratio Test* (MLRT)

Maximum Likelihood Ratio Test merupakan salah satu metode yang digunakan untuk menguji signifikansi parameter dispersi. Pengujian signifikansi parameter dispersi dilakukan untuk menentukan apakah kondisi *equidispersi* terpenuhi atau tidak. Kondisi *equidispersi* dikatakan terpenuhi jika nilai parameter dispersi bernilai 0, dalam hal ini $k = 0$ (Fadhillah, 2011).

Hipotesis Pengujian yang digunakan pada model *Generalized Poisson Regression*:

$H_0 : k = 0$

$H_1 : k \neq 0$

statistik uji:

$$T = -2(\ell_P - \ell_{GP}) = 2(\ell_{GP} - \ell_P) \tag{2.36}$$

ℓ_{GP} adalah nilai log-likelihood pada model *Generalized Poisson Regression* dan ℓ_P adalah nilai log-likehood pada model regresi Poisson (Fadhillah, 2011).

Pada kondisi H_0, statistik uji T mendekati distribusi *Chi-Square* (χ^2) dengan derajat bebas 1. Kriteria pengujiannya adalah tolak H_0 pada taraf signifikansi 0,05 jika $T > {\chi^2}_{(\alpha,1)}$. Penolakan H_0 menunjukkan bahwa *Generalized Poisson Regression* lebih tepat digunakan dibandingkan model regresi Poisson (Fadhillah, 2011).

I. Putus Sekolah

Putus sekolah adalah proses berhentinya siswa secara terpaksa dari suatu lembaga pendidikan tempat dia belajar. Angka putus sekolah mencerminkan persentase anak-anak usia sekolah yang sudah tidak lagi bersekolah atau tidak menamatkan jenjang pendidikan tertentu (Riyadi, dkk., 2015). Angka partisispasi sekolah merupakan proporsi penduduk kelompok usia tertentu yang masih duduk di bangku sekolah (BPS, 2015)

J. Penelitian Terdahulu

Ada beberapa penelitian terdahulu yang juga menerapkan analisis regresi Poisson dan analisis *Generalized Poisson Regression* untuk menyelesaikan suatu permasalahan, diantaranya yaitu:

Fitriana Fadhillah (2011) dalam penelitiannya yang berjudul "Aplikasi Regresi Binomial Negatif dan *Generalized Poisson Regression* dalam mengatasi *Overdispersion* pada Regresi Poisson". Fitriani membandingkan dua model regresi yang merupakan perluasan dari regresi Poisson dalam menangani masalah overdispersi pada kasus data kemiskinan provinsi di Indonesia tahun 2009 untuk mendapatkan faktor-faktor yang mempengaruhi angka kemiskinan provinsi di Indonesia tahun 2009. Hasil analisis dalam penelitian ini menunjukkan bahwa faktor yang berpengaruh terhadap jumlah penduduk miskin adalah jumlah penduduk. Model regresi *Generalized Poisson* memenuhi kriteria kesesuaian model regresi dibandingkan dengan regresi Poisson dan Binomial Negatif.

Ruliana (2015) dalam penelitiannya yang berjudul "Pemodelan Generalized Poisson Regression (GPR) Untuk Mengatasi Pelanggaran Equidispersi Pada Regresi Poisson Kasus Campak Di Kota Semarang" menyatakan bahwa dalam analisis Regresi Poisson seringkali terjadi pelanggaran asumsi *equidispersi* sehingga model regresi Poisson tidak tepat digunakan. Untuk mengatasi masalah tersebut Ruliana menggunakan metode *Generalized Poisson Regression* dalam pemodelan data sehingga didapatkan faktor-faktor yang mempengaruhi jumlah kasus penyakit campak di Kota Semarang tahun 2013.

Adapun penelitian terdahulu yang juga mengangkat kasus anak putus sekolah dalam penilitiannya, yaitu beberapa diantaranya :

Bagus Naufal Fitroni (2013) dalam penelitiannya yang berjudul "Pemodelan Angka Putus Sekolah Usia Wajib Belajar Menggunakan Metode Regresi Spasial di Jawa Timur" menerapkan aspek spasial untuk melihat apakah keragaman karakteristik antar kabupaten/kota di Jawa Timur menentukan kualitas pendidikan pada daerah tersebut, selain itu Fitroni juga ingin mengetahui faktor-faktor yang berpengaruh secara nyata terhadap angka putus sekolah di Jawa Timur. Hasil akhir dari penelitian ini menunjukkan bahwa APS (Angka Putus Sekolah) tingkat SD tidak terdapat dependensi spasial. Hal ini menandakan bahwa pada tingkat SD di setiap kabupaten/kota di Jawa Timur sudah berjalan mandiri dan tidak memiliki ketergantungan dengan daerah yang lainnya. Faktor-faktor yang diperoleh, yaitu rasio guru dan murid, PDRB per kapita, dan rasio penduduk tamatan maksimal SD.

Astari (2013) dalam penelitiannya yang berjudul "Pemodelan Jumlah Anak Putus Sekolah Di Provinsi Bali Dengan Pendekatan *Semi-Parametric Geographically Weighted Poisson Regression*" mengatakan bahwa Faktor-faktor yang mempengaruhi jumlah putus sekolah pada tiap wilayah berbedabeda tergantung pada karakteristik dari masing-masing daerah tersebut, sehingga diperlukan suatu analisis statistika yang memperhitungkan faktor geografis pada tiap pengamatannya. Berdasarkan hasil analisis dan pembahasan yang dilakukan, Astari menyimpulkan bahwa faktor spasial/lokasi tidak berpengaruh terhadap data jumlah anak putus sekolah usia pendidikan dasar di Provinsi Bali tahun 2010. Astari menyatakan bahwa model yang dihasilkan menunjukkan data jumlah anak putus sekolah usia pendidikan dasar di Provinsi Bali tahun 2010 mengandung kasus overdispersi, oleh karena itu model terbaik untuk data jumlah anak putus sekolah usia pendidikan dasar di Provinsi Bali tahun 2010 adalah model regresi yang dapat mengatasi adanya kasus overdispersi, misalnya model regresi Binomial Negatif dan regresi Generalized Poisson. Astari menyatakan, variabel-variabel yang berpengaruh signifikan terhadap jumlah anak putus sekolah usia pendidikan dasar di Provinsi Bali tahun 2010 berdasarkan model regresi Poisson adalah rasio siswa terhadap sekolah pada tiap kecamatan, rasio siswa terhadap guru pada tiap kecamatan, jumlah kepala keluarga dengan pendidikan terakhir ayah SD atau SMP, angka buta huruf, dan rata-rata jumlah anggota keluarga, dan rasio siswa terhadap guru pada tiap kecamatan merupakan variabel yang memiliki pengaruh paling besar.

BAB III

METODOLOGI PENELITIAN

A. Metode Penelitian

Penelitian ini merupakan penerapan model regresi nonlinear, yaitu model regresi Poisson dan model Generalized Poisson Regression pada kasus angka putus sekolah usia wajib belajar di Provinsi Sulawesi Selatan. Penelitian ini dilakukan berdasarkan kajian-kajian teori tentang penelitian-penelitian terdahulu dengan mengumpulkan literarur-literatur yang berkaitan.

B. Sumber Data

Data yang akan digunakan untuk penelitian ini merupakan data sekunder yang diperoleh dari instansi pemerintahan, yaitu Badan Pusat Statistik Provinsi Sulawesi Selatan pada tahun 2015.

C. Waktu dan Tempat Penelitian

Berdasarkan jenis penilitian yang digunakan dalam maka penelitian ini akan dilaksanakan di Perpustakaan Jurusan Matematika Fakultas Matematika dan IPA Universitas Negeri Makassar dengan mengumpulkan literatur-literatur yang berkaitan dan mendukung penelitian ini. Penelitian ini dilaksanakan pada bulan Januari tahun 2017 sampai bulan Juli 2017 dan bertempat di Kota Makassar.

D. Variabel Penelitian

a. Variabel Respon

Pada penelitian ini variabel respon yang digunakan adalah jumlah angka putus sekolah pada usia wajib belajar, yaitu penduduk usia 7-18 tahun.

b. Variabel Prediktor

1. Rasio siswa terhadap terhadap jumlah sekolah

   Rasio siswa terhadap sekolah adalah perbandingan antara jumlah siswa dengan jumlah sekolah.

2. Rasio siswa terhadap guru.

   Perbandingan antara jumlah siswa dengan jumlah guru. Rasio jumlah guru terhadap siswa menggambarkan banyaknya murid yang diajar oleh seorang guru.

3. Persentase penduduk miskin

   Penduduk dengan kondisi kehidupan serba kekurangan sehingga tidak mampu memenuhi kehidupan yang layak bagi kehidupannya.

4. Kepadatan penduduk

   Angka kepadatan penduduk menggambarkan rata-rata jumlah penduduk tiap satu kilometer persegi di suatu wilayah. Semakin besar angka kepadatan penduduk menunjukkan semakin padat penduduk yang mendiami wilayah tersebut.

5. Rata-rata lama sekolah

   Rata-rata lama sekolah menggambarkan tingkat pencapaian setiap penduduk dalam kegiatan bersekolah. Semakin tinggi angka lamanya bersekolah semakin tinggijenjang pendidikan yang telah dicapai penduduk.

E. Prosedur Kerja

Berikut ini merupakan prosedur kerja yang akan dilakukan dalam analisis data untuk mencapai tujuan penelitian dan dijelaskan oleh Gambar 3.1:

1. Mengumpulkan data dan asumsikan berdistribusi Poisson
2. Mendeteksi hubungan kolonieritas antara variabel prediktor

   Ada tidaknya multikolinieritas dapat diketahui dengan melihat nilai VIF (*Variance Inflation Factor*) dan *Tollerance* yang diberikan oleh persamaan (2.25) dan persamaan (2.26).
3. Menaksir parameter model regresi Poisson dengan menggunakan metode *Maximum Likelihood Estimation*.

   Penaksiran parameter menggunakan metode *Maximum Likelihood Estimation* dapat dilakukan denga langkah-langkah :

   a. Mengambil n buah sampel
   b. Membentuk fungsi likelihood berdasarkan persamaan (2.12) sehingga diperoleh fungsi *likelihood* untuk model regresi Poisson sebagaimana persamaan (2.14)
   c. Mengambil bentuk log dari fungsi *likelihood* yang diperoleh bentuk dari fungsi log-*likelihood*nya seperti yang ditunjukkan pada persaaman (2.15)
   d. Mendiferensialkan fungsi log-*likelihood*nya terhadap masing-masing parameter $\boldsymbol{\beta}$ dan disamakan denga nol. Sehingga akan diperoleh taksiran parameter atas penyelesaian persamaan (2.15) seperti persamaan (2.16)
4. Melakukan pengujian parameter model regresi Poisson

Pengujian parameter pada model regresi Poisson dilakukan mennggunakan Uji Serentak dan Uji Parsial yang statistik ujinya masing-masing ditunjukkan oleh persamaan (2.17) dan (2.18).

5. Mendapatkan model regresi Poisson
6. Uji asumsi *equdispersi*

Uji asumsi *equidispersi* dapat dilakukan dengan melihat nilai mean dan nilai varians dari variabel Y, dikatakan memenuhi asumsi jika nilai mean dan nilai varians memiliki nilai yang sama. Jika nilai varians lebih besar dari nilai mean maka data dikatakan mengalami *overdispersi*, dan jika nilai varians lebih kecil dari nilai mean maka data dikatakan mengalami *Underdispersi*, sehingga pemodelan sebaiknya dilakukan menggunakan metode *Generalized Poisson Regression* .

Jika uji asumsi *equidispersi* terpenuhi maka langkah kerja akan dilanjutkan pada langkah kerja ke-11. Jika asumsi *equidispersi* tidak terpenuhi maka langkah kerja akan dilanjutkan pada langkah ke-7.

7. Menaksir Parameter model GPR dengan *Maximum Likelihood Estimation*

Penaksiran parameter model GPR dilakukan dengan metode *Maximum Likelihood Estimation* (MLE). Pada umumnya penaksiran parameter model *Generalized Poisson Regression* memiliki langkah yang sama dengan penaksiran parameter pada model regresi Poisson.

8. Menguji signifikansi model GPR

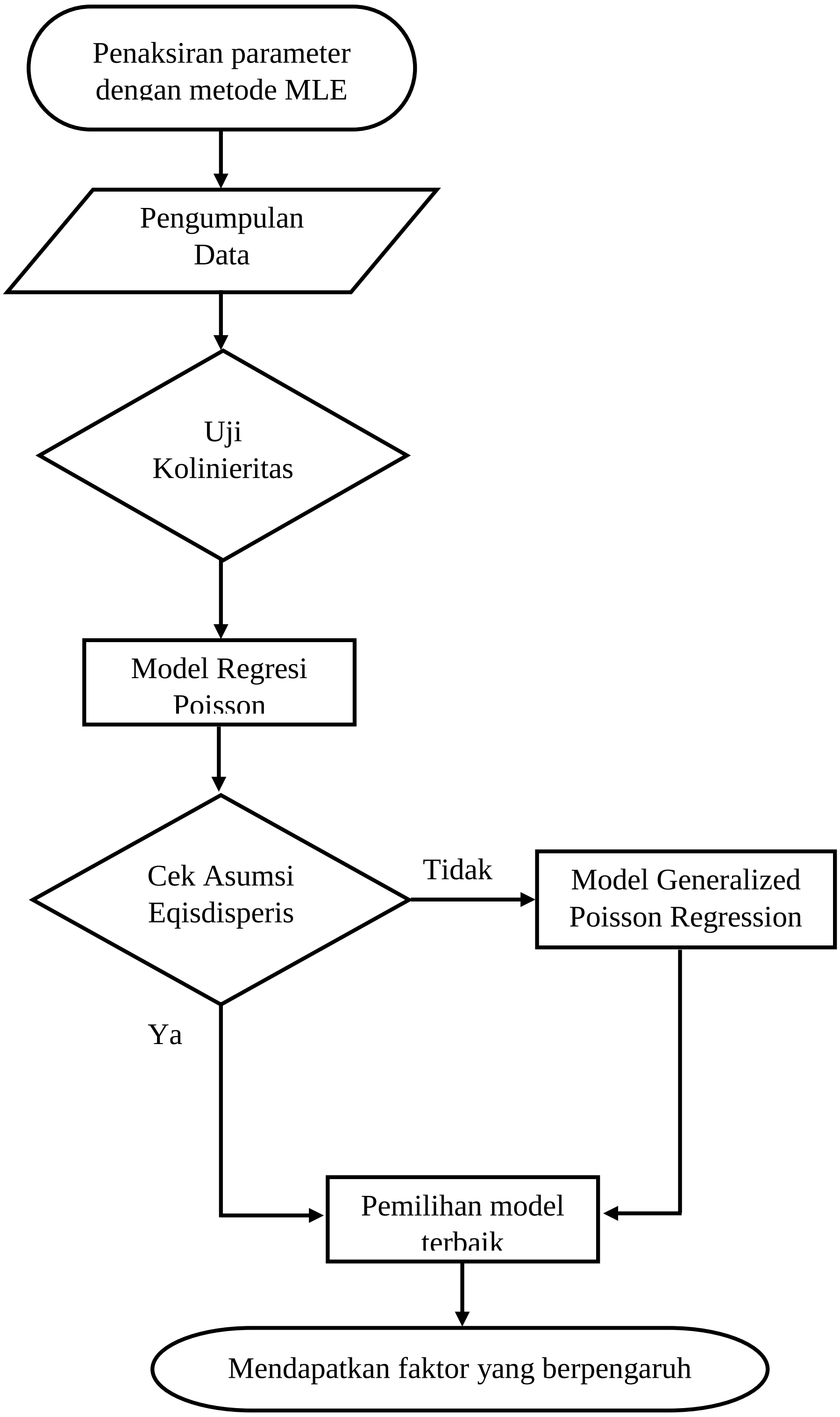
Pengujian parameter untuk model *Generalized Poisson Regression* dilakukan sama dengan pengujian parameter pada regresi Poisson, yaitu dengan menggunakan metode *Maximum Likelihood Ratio Test* (MLRT) sebagaimana persamaan (2.37) untuk pengujian parameter secara simultan dan menggunakan statistik uji Wald sebagaimana persamaan (2.36) untuk pengujian parameter secara parsial.

9. Mendapatkan model GPR
10. Memilih model terbaik

    Pemilihan model terbaik dilakukan dengan melihat nilai AIC. Model yang memiliki nilai AIC terkecil merupakan model terbaik.

11. Mendapatkan faktor yang mempengaruhi angka putus sekolah sekolah pada usia wajb belajar.

F. Bagan



Gambar 3.1 prosedur Kerja

BAB IV

HASIL DAN PEMBAHASAN

A. Hasil

1. Eksplorasi Data

Data yang digunakan pada penelitian ini merupakan data sekunder yang diperoleh dari Badan Pusat Statistik Provinsi Sulawesi Selatan. Data yang digunakan merupakan data angka putus sekolah usia wajib belajar yang diasumsikan mengikuti distribusi Poisson beserta faktor-faktor yang diduga berpengaruh terhadap pertumbuhan angka putus sekolah di Provinsi Sulawesi Selatan tahun 2015. Data ini diperoleh berdasarkan hasil Survey Sosial Ekonomi Nasional (SUSENAS) tahun 2016 dan data Sulawesi Selatan dalam Angka tahun 2016. Data yang akan digunakan dalam penelitian ini dapat dilihat pada Tabel 4.1.

Tabel 4.1 Data Penelitian

Kode.	Kabupaten	Y	X1	X2	X3	X4	X5
1	Selayar	1473	124.54	10.70	16900	144	7.16
2	Bulukumba	8879	168.26	13.03	33360	355	6.68
3	Bantaeng	4941	166.59	10.55	17550	463	6.16
4	Jeneponto	9431	169.87	12.39	53870	394	5.64
5	Takalar	6205	176.22	11.82	27120	506	6.57
6	Gowa	15852	218.30	15.29	59470	384	7.24
7	Sinjai	4631	155.26	10.65	21990	290	7.05
8	Maros	8134	172.25	13.61	40080	210	7.19
9	Pangkep	11418	149.62	12.01	53850	291	7.32
10	Barru	2992	130.00	10.61	16100	146	7.60
11	Bone	15995	154.50	12.85	75010	163	6.55
12	Soppeng	3331	120.23	11.49	18880	166	7.05
13	Wajo	12331	121.93	10.93	30080	157	6.37
14	Sidrap	5804	158.36	12.53	16030	154	7.32
15	Pinrang	8841	177.57	12.73	30510	187	7.47
16	Enrekang	3191	153.19	11.03	27600	112	8.05

17	Luwu	7144	189.78	12.96	48640	117	7.74
18	Tana Toraja	2421	166.10	13.07	28590	111	7.91
22	Luwu Utara	6424	185.63	13.53	41890	40	7.38
25	Luwu Timur	4646	250.00	16.09	19670	40	7.87
26	Toraja Utara	3295	214.23	15.08	34370	196	7.71
71	Makassar	27765	308.31	17.94	63240	8246	10.77
72	Parepare	2470	208.60	11.93	8410	1396	10.01
73	Palopo	2206	287.72	13.84	14510	682	10.25

Sumber: BPS Provinsi Sulawesi Selatan

Keterangan:

Y = Jumlah angka putus sekolah.

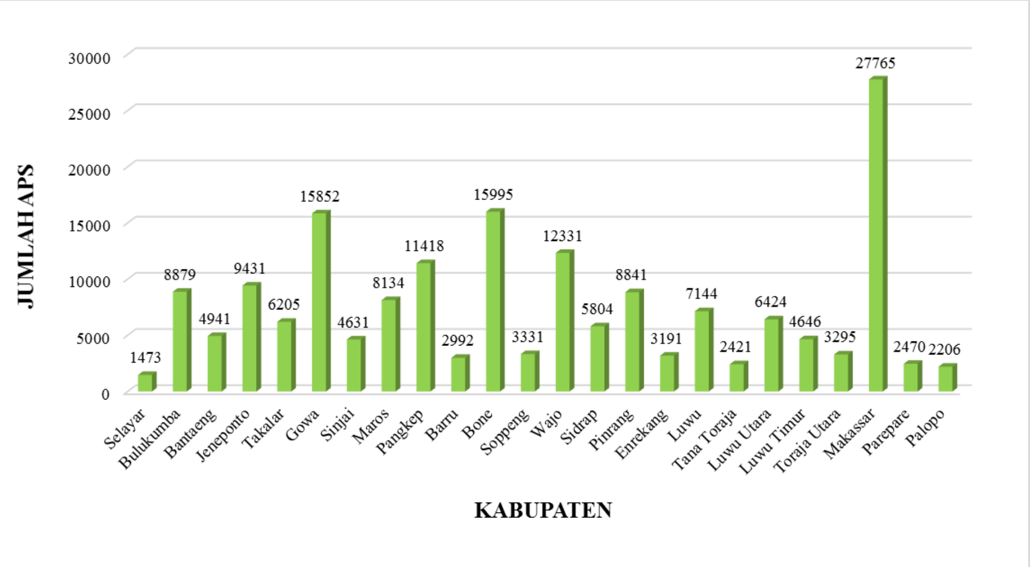
X1 = Rasio siswa terhadap sekolah

X2 = Rasio siswa terhadap guru

X3 = Jumlah penduduk miskin (ribu)

X4 = Kepadatan penduduk (orang/Km^2)

X5 = Rata-rata lama bersekolah.



Gambar 4.1 Jumlah Angka Putus Sekolah pada tiap Kabupaten/Kota

Gambar 4.1 memperlihatkan grafik pertumbuhan angka putus sekolah di Provinsi Sulawesi Selatan pada tahun 2015. Pada gambar dapat dilihat angka putus sekolah paling tinggi terdapat di Kota Makassar, yaitu sebesar 27.765 orang dan angka putus sekolah palin rendah terdapat di Kabupaten Selayar, yaitu sebesar 1.473 orang. Dalam penelitian ini ingin diketahui seberapa besar pengaruh faktor-faktor yang diteliti terhadap pertumbuhan angka putus sekolah pada usia wajib belajar yang terjadi di Sulawesi Selatan pada tahun 2015.

2. Penaksiran Parameter Model Generalized Poisson Regression menggunakan Metode Maximum Likelihood Estimation

Sebagaimana yang telah dijelaskan pada BAB II, penaksiran parameter pada model Generalized Poisson Regression pada umumnya memiliki langkah yang sama dengan penaksiran parameter pada model regresi Poisson. Penaksiran parameter β dan k dapat dilakukan denagn metode *Maximum Likelihood Estimation*, sebagaimana langkah-langkah berikut:

1) Membentuk fungsi Likelihood. Berikut merupakan fungsi likelihood untuk model *Generalized Poisson Regression* sebagaimana yang telah ditunjukkan pada persamaan (2.30)

$$L(\boldsymbol{\beta},k,y) = \prod_{i=1}^{n} f(y_i,\boldsymbol{\beta},k)$$

$$= \prod_{i=1}^{n}\left(\frac{\mu_i}{1+k\mu_i}\right)^{y_i}\frac{(1+ky_i)^{y_i-1}}{y_i!}\exp\left(-\frac{\mu_i(1+ky_i)}{(1+k\mu_i)}\right)$$

$$= \prod_{i=1}^{n}\left\{\left(\frac{\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}\right)}{1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}\right)}\right)^{y_i}\frac{(1+ky_i)^{y_i-1}}{y_i!}\right.$$

$$\left.\exp\left(-\frac{\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}\right)(1+ky_i)}{\left(1+k\exp\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}\right)\right)}\right)\right\}$$

2) Membentuk fungsi log-likelihood dari fungsi likelihood yang telah diperoleh

$$\ln L(\boldsymbol{\beta}, k, y) = \ln \prod_{i=1}^{n} f(y_i, \boldsymbol{\beta}, k)$$

$$= \ln \prod_{i=1}^{n} \left\{ \left(\frac{\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}\right)}{1 + k \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}\right)} \right)^{y_i} \frac{(1 + k y_i)^{y_i - 1}}{y_i!} \right.$$

$$\left. \exp\left(- \frac{\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}\right)(1 + k y_i)}{\left(1 + k \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}\right)\right)} \right) \right\}$$

$$= \sum_{i=1}^{n} \ln \left\{ \left(\frac{\exp(x_i^T \beta)}{1 + k \exp(x_i^T \beta)} \right)^{y_i} \frac{(1 + k y_i)^{y_i - 1}}{y_i!} \right.$$

$$\left. \exp\left(- \frac{\exp(x_i^T \beta)(1 + k y_i)}{(1 + k \exp(x_i^T \beta))} \right) \right\}$$

$$= \sum_{i=1}^{n} \left\{ \ln \left(\frac{\exp(x_i^T \beta)}{1 + k \exp(x_i^T \beta)} \right)^{y_i} + \ln \frac{(1 + k y_i)^{y_i - 1}}{y_i!} + \right.$$

$$\left. \ln\left(\exp\left(- \frac{\exp(x_i^T \beta)(1 + k y_i)}{(1 + k \exp(x_i^T \beta))} \right) \right) \right\}$$

$$= \sum_{i=1}^{n} \left\{ y_i \ln \left(\frac{\exp(x_i^T \beta)}{1 + k \exp(x_i^T \beta)} \right) + (y_i - 1) \ln(1 + k y_i) - \right.$$

$$\left. \frac{\exp(x_i^T \beta)(1 + k y_i)}{(1 + k \exp(x_i^T \beta))} - \ln y_i! \right\}$$

$$= \sum_{i-1}^{n} \{ y_i \ln(\exp(x_i^T \beta)) - y_i \ln(1 + k \exp(x_i^T \beta)) +$$

$$\left. (y_i - 1) \ln(1 + k y_i) - \frac{\exp(x_i^T \beta)(1 + k y_i)}{(1 + k \exp(x_i^T \beta))} - \ln y_i! \right\}$$

$$= \sum_{i=1}^{n} \{ y_i (x_i^T \beta) - y_i \ln(1 + k \exp(x_i^T \beta)) +$$

$$\left. (y_i - 1) \ln(1 + k y_i) - \frac{\exp(x_i^T \beta)(1 + k y_i)}{(1 + k \exp(x_i^T \beta))} - \ln y_i! \right\}$$

3) Mendiferensialkan persamaan log-likelihood yang telah diperoleh

a. Mendiferensilakan fungsi log-likelihood pada persamaan (2.31) terhadap $\boldsymbol{\beta}_0$

$$\frac{\partial \ln L(\beta,k)}{\partial \beta_0} = 0$$

$$= \sum_{i=0}^{n} y_i - y_i \frac{k\exp(x_i^T\beta)}{1+k\exp(x_i^T\beta)} - \frac{(1+ky_i)\exp(x_i^T\beta)(1+k\exp(x_i^T\beta)) - k\exp(x_i^T\beta)(1+ky_i)\exp(x_i^T\beta)}{\left(1+k\exp(x_i^T\beta)\right)^2}$$

$$= \sum_{i=0}^{n} y_i - y_i \frac{k\exp(x_i^T\beta)}{1+k\exp(x_i^T\beta)} - \frac{(1+ky_i)\exp(x_i^T\beta)\left[(1+k\exp(x_i^T\beta)) - k\exp(x_i^T\beta)\right]}{\left(1+k\exp(x_i^T\beta)\right)^2}$$

$$= \sum_{i=0}^{n} \left\{ y_i \left(1 - \frac{k\exp(x_i^T\beta)}{1+k\exp(x_i^T\beta)}\right) - \frac{(1+ky_i)\exp(x_i^T\beta)}{\left(1+k\exp(x_i^T\beta)\right)^2} \right\}$$

$$= \sum_{i=0}^{n} \left\{ y_i \left(\frac{1+k\exp(x_i^T\beta) - k\exp(x_i^T\beta)}{1+k\exp(x_i^T\beta)}\right) - \frac{(1+ky_i)\exp(x_i^T\beta)}{\left(1+k\exp(x_i^T\beta)\right)^2} \right\}$$

$$= \sum_{i=0}^{n} \left\{ \frac{y_i}{1+k\exp(x_i^T\beta)} - \frac{(1+ky_i)\exp(x_i^T\beta)}{\left(1+k\exp(x_i^T\beta)\right)^2} \right\}$$

$$= \sum_{i=0}^{n} \left(\frac{y_i(1+k\exp(x_i^T\beta)) - (1+ky_i)\exp(x_i^T\beta)}{\left(1+k\exp(x_i^T\beta)\right)^2} \right)$$

$$= \sum_{i=0}^{n} \left(\frac{y_i + ky_i\exp(x_i^T\beta) - \exp(x_i^T\beta) - ky_i\exp(x_i^T\beta)}{\left(1+k\exp(x_i^T\beta)\right)^2} \right)$$

$$= \sum_{i=1}^{n} \left(\frac{y_i - \exp(x_i^T\beta)}{\left(1+k\exp(x_i^T\beta)\right)^2} \right)$$

$$= \sum_{i=1}^{n} \left(\frac{y_i - \mu_i}{(1+k\mu_i)^2} \right) = 0$$

b. Mendiferensilakan fungsi log-likelihood pada persamaan (2.31) terhadap $\boldsymbol{\beta}_j$

$$\frac{\partial \ln L(\beta,k)}{\partial \beta_j} = 0$$

$$= \sum_{i=1}^{n} \left\{ y_i x_{ij} - y_i x_{ij} \frac{k \exp(x_i^T \beta)}{1+k \exp(x_i^T \beta)} - \right.$$

$$\left. \frac{(1+ky_i)x_{ij} \exp(x_i^T \beta)(1+k \exp(x_i^T \beta)) - k x_{ij} \exp(x_i^T \beta)(1+ky_i) \exp(x_i^T \beta)}{\left(1+k \exp(x_i^T \beta)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ y_i x_{ij} - y_i x_{ij} \frac{k \exp(x_i^T \beta)}{1+k \exp(x_i^T \beta)} - \right.$$

$$\left. \frac{(1+ky_i)x_{ij} \exp(x_i^T \beta)[1+k \exp(x_i^T \beta) - k \exp(x_i^T \beta)]}{\left(1+k \exp(x_i^T \beta)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ y_i x_{ij} \left(1 - \frac{k \exp(x_i^T \beta)}{1+k \exp(x_i^T \beta)}\right) - \frac{(1+ky_i)x_{ij} \exp(x_i^T \beta)}{\left(1+k \exp(x_i^T \beta)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ y_i x_{ij} \left(\frac{1+k(x_i^T \beta) - k \exp(x_i^T \beta)}{1+k \exp(x_i^T \beta)}\right) - \frac{(1+ky_i)x_{ij} \exp(x_i^T \beta)}{\left(1+k \exp(x_i^T \beta)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ \frac{y_i x_{ij}}{1+k \exp(x_i^T \beta)} - \frac{(1+ky_i)x_{ij} \exp(x_i^T \beta)}{\left(1+k \exp(x_i^T \beta)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ \frac{y_i x_{ij}(1+k \exp(x_i^T \beta)) - (1+ky_i)x_{ij} \exp(x_i^T \beta)}{\left(1+k \exp(x_i^T \beta)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left(\frac{y_i x_{ij} + y_i x_{ij} k \exp(x_i^T \beta) - x_{ij} \exp(x_i^T \beta) - k y_i x_{ij} \exp(x_i^T \beta)}{\left(1+k \exp(x_i^T \beta)\right)^2} \right)$$

$$= \sum_{i=1}^{n} \left(\frac{y_i x_{ij} - x_{ij} \exp(x_i^T \beta)}{\left(1+k \exp(x_i^T \beta)\right)^2} \right)$$

$$= \sum_{i=1}^{n} \left(\frac{x_{ij}(y_i - \exp(x_i^T \beta))}{\left(1+k \exp(x_i^T \beta)\right)^2} \right)$$

$$= \sum_{i=1}^{n} \frac{x_{ij}(y_i - \mu_i)}{(1+k\mu_i)^2} \quad ; j = 1, 2, \cdots, p$$

$$= 0$$

c. Mendiferensilakan fungsi log-likelihood pada persamaan (2.31) terhadap k

$$\frac{\partial \ln L(\beta,k)}{\partial k} = 0$$

$$= \sum_{i=1}^{n} \left\{ -\frac{y_i \exp(x_i^T \beta)}{1+k\exp(x_i^T \beta)} + \frac{y_i(y_i-1)}{1+ky_i} - \right.$$

$$\left. \frac{y_i \exp(x_i^T \beta)(1+k\exp(x_i^T \beta)) - \exp(x_i^T \beta)\exp(x_i^T \beta)(1+ky_i)}{\left(1+k\exp(x_i^T \beta)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ -\frac{y_i \exp(x_i^T \beta)}{1+k\exp(x_i^T \beta)} + \frac{y_i(y_i-1)}{1+ky_i} - \right.$$

$$\left. \frac{y_i \exp(x_i^T \beta) + ky_i \exp(x_i^T \beta) - \exp(x_i^T \beta)^2 - ky_i \exp(x_i^T \beta)}{\left(1+k\exp(x_i^T \beta)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ -\frac{y_i \exp(x_i^T \beta)}{1+k\exp(x_i^T \beta)} + \frac{y_i(y_i-1)}{1+ky_i} - \frac{y_i \exp(x_i^T \beta) - \exp(x_i^T \beta)^2}{\left(1+k\exp(x_i^T \beta)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ -\frac{y_i \exp(x_i^T \beta)}{1+k\exp(x_i^T \beta)} + \frac{y_i(y_i-1)}{1+ky_i} - \frac{\exp(x_i^T \beta)(y_i - \exp(x_i^T \beta))}{\left(1+k\exp(x_i^T \beta)\right)^2} \right\}$$

$$= \sum_{i=1}^{n} \left\{ -\frac{y_i \mu_i}{1+k\mu_i} + \frac{y_i(y_i-1)}{1+ky_i} - \frac{\mu_i(y_i-\mu_i)}{(1+k\mu_i)^2} \right\}$$

$$= 0$$

Penaksiran parameter $\boldsymbol{\beta}$ secara manual mengalami kesulitan karena hasilnya tidak eksak sehingga dalam penentuannya dilakukan menggunakan bantuan *software* statistika, yaitu SPSS.23. Model Generalized Poisson Regression memiliki benuk persamaan yang sama dengan model regresi Poisson sebagaimana persamaan (2.13).

3. Pemodelan Angka Putus Sekolah di Sulawesi Selatan

1) Uji Kolinieritas

Salah satu syarat yang harus dipenuhi dalam analisis regresi adalah tidak terjadi kasus multikolinieritas antar variabel prediktor. Oleh karena itu, dilakukan uji kolinieritas pada data yang diamati. Masalah multikolinieritas dapat dideteksi dengan melihat nilai *Tollerance* atau nilai VIF (*Variance Inflation Factor*) data yang diamati. Dengan bantuan software SPSS.23 maka diperoleh nilai *Tollerance* dan nilai VIF sebagaimana diperlihatkan pada Tabel 4.2. Hasil analisis selengkapanya dapat dilihat pada lampiran 2.

Tabel 4.2 Nilai VIF dan nilai Tolerance setiap variabel prediktor

Variabel	Collienarity Statistic	
	Tollerance	VIF
Rasio Siswa Sekolah	0.168	5.968
Rasio Siswa Guru	0.206	4.846
Jumlah Penduduk Miskin	0.528	1.895
Kepadatan Penduduk	0.454	2.200
Rata-rata Lama Bersekolah	0.308	3.251

Berdasarkan Tabel 4.2 maka dapat dijelaskan bahwa tidak terjadi kasus multikolonieritas pada data yang diamati, yang berarti tidak ada keterkaitan hubungan antara tiap-tiap variabel predikor ditunjukkan oleh nlai VIF untuk tiap-tiap variabel yang lebih kecil dari 10 dan nilai *Tollerance* yang lebih besar dari 0,1. Oleh karena tidak tejadi multikolinieritas pada data yang diamati maka analisis data menggunakan model regresi dapat dilanjutkan.

2) Pemodelan

a. Model Regresi Poisson

Bentuk umum dari model regresi Poisson adalah sebgaimana pada persamaan (2.13), yaitu:

$$y_i = \mu_i + \varepsilon_i$$

$$y_i = \exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi}) + \varepsilon_i$$

Langkah selanjutnya yang akan dilakukan adalah memodelkan faktor-faktor yang mempengaruhi jumlah putus sekolah. Untuk melakukan pomedelan maka terlebih dahulu akan dicari nilai taksiran parameter $\beta_1, \beta_2, \beta_3, \beta_4,$ dan β_5. Penaksiran parameter $\boldsymbol{\beta}$ dicari menggunakan bantuan software SPSS.23. Hasil penaksiran untuk masing-masing parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.3. Hasil analisis selengkapnya dapat dilihat pada lampiran 3.

Tabel 4.3 Hasil penaksiran parameter model regresi Poisson

Parameter	Taksiran $\boldsymbol{\beta}$	Std. Error
(Konstan)	8.8773316429	0.0319655225
X1	-0.0021822726	0.0001561293
X2	0.0869555778	0.0035345523
X3	0.0000202693	0.0000001774
X4	0.0001877907	0.0000020579
X5	-0.2190109654	0.0039372724

Setelah mendapatkan hasil taksiran parameter maka selanjutnya akan dilakukan pengujian parameter utnuk mendapatkan variabel prediktor yang berpengaruh secara nyata terhadap variabel respon. Pengujian parameter dilakukan menggunakan uji Wald sebagaimana persamaan (2.18).

$$W_1 = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{-0.0021822726}{0.0001561293}\right)^2 = 195.3660298$$

$$W_2 = \left(\frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}\right)^2 = \left(\frac{0.0869555778}{0.0035345523}\right)^2 = 605.2378256$$

$$W_3 = \left(\frac{\hat{\beta}_3}{SE(\hat{\beta}_3)}\right)^2 = \left(\frac{0.0000202693}{0.0000001774}\right)^2 = 13057.71302$$

$$W_4 = \left(\frac{\hat{\beta}_4}{SE(\hat{\beta}_4)}\right)^2 = \left(\frac{0.0001877907}{0.0000020579}\right)^2 = 8327.4501$$

$$W_5 = \left(\frac{\hat{\beta}_5}{SE(\hat{\beta}_5)}\right)^2 = \left(\frac{-0.2190109654}{0.0039372724}\right)^2 = 3094.145937$$

Kriteria pengujian untuk uji Wald adalah tolak H_0 jika niai dari uji Wald lebih besar dari nilai ${\chi^2}_{(\alpha,1)}$. Berdasarkan hasil pengujian parameter yang telah dilakukan maka dapat dijelaskan bahwa setiap variabel prediktor yang diamati memiliki pengaruh secara nyata terhadap variabel respon. Dengan demikian model angka putus sekolah menggunakan regresi Poisson dapat dituliskan sebagai berikut:

$$y_i = \exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi}) + \varepsilon_i$$

$$Y = \exp(8.8773316429 - 0.0021822726X_1 + 0.0869555778X_2 + 0.0000202693X_3 + 0.0001877907X_4 - 0.2190109654X_5) + \varepsilon_i$$

b. Uji Asumsi Equidispersi

Terdapat asumsi yang harus dipenuhi dalam regresi Poisson yaitu niai mean dan nilai varians dari variabel respon harus memiliki nilai yang sama. Nilai mean dan varians dari variabel respon diperlihatkan pada Tabel 4.3.

Tabel 4.4 Statistika Deskriptif

Variabel Respon	N	Mean	Std. Deviation	Variance
Angka Putus Sekolah	24	7492.50	5961.008	35533612

$$\bar{x} = \frac{\sum_{i=1}^{24} x_i}{n} = \frac{179820}{24} = 7492{,}5$$

$$s^2 = \frac{\sum_{i=1}^{24}(x_i - \bar{x})^2}{n-1} = \frac{817273076}{23} = 35533612$$

Berdasarkan nilai mean dan varians variabel respon diatas maka dapat disimpulkan bahwa model regeresi poisson biasa tidak dapat digunakan untuk melanjutkan pemodelan angka putus sekolah karena tidak memenuhi asumsi *equidispersi*. Dapat dilihat juga bahwa nilai varians dari variabel respon memiliki nilai yang lebih besar daripada nilai mean dari variabel respon yang menandakan bahwa data yang diamati mengalami *Overdispersi*.

c. Model *Generalized Poisson Regression*

Pada dasarnya langkah untuk mendapatkan model *Generalized Poisson Regression* sama dengan model Regresi Poisson biasa namun pada model *Generalized Poisson Regression* dilakukan dengan melibatkan satu atau beberapa variabel prediktor. Degan demikian akan didapatkan beberapa model dan akan dipilih berdasarkan nilai AICnya.

1. Model yang terbentuk dengan melibatkan satu variabel prediktor

Berikut ini beberapa model yang akan terbentuk jika hanya melibatkan satu variabel prediktor saja, dengan memasukkan $X_1, X_2, X_3, X_4,$ dan X_5 pada model secara bergantian. Hasil analisis selenkapnya untuk model dengan satu variabel prediktor dapat dilihat pada lampiran 4.

a. Model dengan hanya X_1 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.5 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.5 Taksiran parameter untuk model dengan hanya melibatkan X_1 sebagai variabel prediktor

Parameter	$\boldsymbol{\beta}$	Std. Error
(Konstan)	7.97081210353	0.00880290690
X1	0.00509843188	0.00004388095

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{0.00509843188}{0.00004388095}\right)^2 = 13499{,}6111$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel X_1 memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(7.97081210353 + 0.00509843188X_1) + \varepsilon_i$$

b. Model dengan hanya X_2 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.6 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.6 Taksiran parameter untuk model dengan hanya melibatkan $\boldsymbol{X_2}$ sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	6.36309	0.01502
X2	0.19462	0.00109

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald

$$W = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{0.19462}{0.00109}\right)^2 = 31638{,}229$$

Berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel X_2 memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(6.36309 + 0.19462X_1) + \varepsilon_i$$

c. Model dengan hanya X_3 sebagai variable prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.7 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.7 Taksiran parameter untuk model dengan hanya melibatkan $\boldsymbol{X_3}$ sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	7.7426289021	0.0058347211
X3	0.0000307115	0.0000001215

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W = \left(\frac{\hat{\beta}_3}{SE(\hat{\beta}_3)}\right)^2 = \left(\frac{0.0000307115}{0.0000001215}\right)^2 = 63934{,}189$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel X_2 memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(7{,}7426289021 + 0.0000307115X_3) + \varepsilon_i$$

d. Model dengan hanya X_4 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.8 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.8 Taksiran parameter untuk model dengan hanya melibatkan X_4 sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	8.7487804379	0.0026644268
X4	0.0001764386	0.0000008275

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W = \left(\frac{\hat{\beta}_4}{SE(\hat{\beta}_4)}\right)^2 = \left(\frac{0.0001764386}{0.0000008275}\right)^2 = 45463.5166$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel X_4 memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(8{,}7487804379 + 0{,}0001764386X_4) + \varepsilon_i$$

e. Model dengan hanya melibatkan X_5 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.9 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.9 Taksiran parameter untuk model dengan hanya melibatkan $\boldsymbol{X_5}$ sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	8.3449793698…	0.0143197118…
X5	0.0758622572…	0.0018435520…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W = \left(\frac{\hat{\beta}_5}{SE(\hat{\beta}_5)}\right)^2 = \left(\frac{0.0758622572}{0.0018435520}\right)^2 = 1693.326561$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel X_5 memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(8.3449793698 + 0.0758622572X_5) + \varepsilon_i$$

2. Model dengan melibatkan dua variabel prediktor

Berikut ini beberapa model yang akan terbentuk jika hanya melibatkan dua variabel prediktor saja, pada model. Hasil analisis selenkapnya untuk model dengan satu variabel prediktor dapat dilihat pada lampiran 5

a. Model dengan hanya X_1 & X_2 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.10 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.10 Taksiran parameter untuk model dengan hanya melibatkan X_1 & X_2 sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	5.57771…	0.01963…
X1	-0.00703…	0.00011…
X2	0.35375…	0.00265…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_1 = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{-0.00703}{0.00011}\right)^2 = 4456.016$$

$$W_2 = \left(\frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}\right)^2 = \left(\frac{0.35375}{0.00265}\right)^2 = 17829.98$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel X_1 & X_2 memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(5.57771 - 0.00703X_1 + 0.35375X_2) + \varepsilon_i$$

b. Model dengan hanya X_1 & X_3 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.11 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.11 Taksiran parameter untuk model dengan hanya melibatkan X_1 & X_3 sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	7.2717352...	0.0094184297
X1	0.0028661...	0.0000446092…
X3	0.0000290...	0.0000001251…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_1 = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{0.0028661}{0.0000446092}\right)^2 = 4128.128$$

$$W_3 = \left(\frac{\hat{\beta}_3}{SE(\hat{\beta}_3)}\right)^2 = \left(\frac{0.0000290}{0.0000001251}\right)^2 = 53973.739$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel X_1 & X_3 memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(7.2717352 + 0.0028661X_1 + 0.0000290X_3) + \varepsilon_i$$

c. Model dengan hanya X_1 & X_4 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.12 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.12 Taksiran parameter untuk model dengan hanya melibatkan X_1 & X_4 sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	9.0522535181…	0.0119201465…
X1	-0.0018003513…	0.0000694599…
X4	0.0002074042…	0.0000014609…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_1 = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{-0.0018003513}{0.0000694599}\right)^2 = 671.809$$

$$W_4 = \left(\frac{\hat{\beta}_4}{SE(\hat{\beta}_4)}\right)^2 = \left(\frac{0.0002074042}{0.0000014609}\right)^2 = 20154.987$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel X_1 & X_4 memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(9.052253 - 0.001800X_1 + 0.000207X_4) + \varepsilon_i$$

d. Model dengan hanya X_1 & X_5 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.13 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.13 Taksiran parameter untuk model dengan hanya melibatkan X_1 & X_5 sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	8.7234642008…	0.0138518356…

X1	0.0093005244…	0.0000747214…
X5	-0.2024013940…	0.0028500180…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_1 = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{0.0093005244}{0.0000747214}\right)^2 = 15492.639$$

$$W_5 = \left(\frac{\hat{\beta}_5}{SE(\hat{\beta}_5)}\right)^2 = \left(\frac{-0.2024013940}{0.0028500180}\right)^2 = 5043.497$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel X_1 & X_5 memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(8.7234642 + 0.0093005X_1 - 0.2024013X_5) + \varepsilon_i$$

e. Model dengan hanya X_2 & X_3 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.14 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.14 Taksiran parameter untuk model dengan hanya melibatkan X_2 & X_3 sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	6.7937184058…	0.0153109354…
X2	0.0832837908…	0.0012379991…
X3	0.0000268875…	0.0000001368…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_2 = \left(\frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}\right)^2 = \left(\frac{0.0832837908}{0.0012379991}\right)^2 = 4525.643$$

$$W_3 = \left(\frac{\hat{\beta}_3}{SE(\hat{\beta}_3)}\right)^2 = \left(\frac{0.0000268875}{0.0000001368}\right)^2 = 38641.141$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel X_2 & X_3 memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(6.7937184 + 0.0832837X_2 + 0.0000268X_3) + \varepsilon_i$$

f. Model dengan hanya X_2 & X_4 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.15 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.15 Taksiran parameter untuk model dengan hanya melibatkan X_2 & X_4 sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	7.55877986…	0.02024483…
X2	0.09538508…	0.00159268…
X4	0.00011412…	0.00000131…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_2 = \left(\frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}\right)^2 = \left(\frac{0.09538508}{0.00159268}\right)^2 = 3586.791$$

$$W_4 = \left(\frac{\hat{\beta}_4}{SE(\hat{\beta}_4)}\right)^2 = \left(\frac{0.00011412}{0.00000131}\right)^2 = 7608.903$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel X_2 & X_4 memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(7.558779 - 0.095385X_2 + 0.000114X_4) + \varepsilon_i$$

g. Model dengan hanya X_2 & X_5 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.16 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.16 Taksiran parameter untuk model dengan hanya melibatkan X_2 & X_5 sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	6.7161…	0.0165…
X2	0.2553…	0.0015…
X5	-0.1519…	0.0023…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_2 = \left(\frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}\right)^2 = \left(\frac{0.2553}{0.0015}\right)^2 = 30604.120$$

$$W_5 = \left(\frac{\hat{\beta}_5}{SE(\hat{\beta}_5)}\right)^2 = \left(\frac{-0.1519}{0.0023}\right)^2 = 4242.499$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel X_2 & X_5 memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(6.7161 + 0.2553X_2 - -0.1519X_5) + \varepsilon_i$$

h. Model dengan hanya X_3 & X_4 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.17 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.17 Taksiran parameter untuk model dengan hanya melibatkan X_3 & X_4 sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	7.8177454557…	0.0059222367…
X3	0.0000265246…	0.0000001329…
X4	0.0000892477…	0.0000008942…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_3 = \left(\frac{\hat{\beta}_3}{SE(\hat{\beta}_3)}\right)^2 = \left(\frac{0.0000265246}{0.0000001329}\right)^2 = 39816.788$$

$$W_4 = \left(\frac{\hat{\beta}_4}{SE(\hat{\beta}_4)}\right)^2 = \left(\frac{0.0000892477}{0.0000008942}\right)^2 = 9962.612$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel X_3 & X_4 memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(7.8177454 + 0.0000265X_3 + 0.0000892X_4) + \varepsilon_i$$

i. Model dengan hanya X_3 & X_5 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.18 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.18 Taksiran parameter untuk model dengan hanya melibatkan X_3 & X_5 sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	7.0732874239	0.0138521996
X3	0.0000307319…	0.0000001210…
X5	0.0882285118…	0.0016561044…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_3 = \left(\frac{\hat{\beta}_3}{SE(\hat{\beta}_3)}\right)^2 = \left(\frac{0.0000307319}{0.0000001210}\right)^2 = 64559.445$$

$$W_5 = \left(\frac{\hat{\beta}_5}{SE(\hat{\beta}_5)}\right)^2 = \left(\frac{0.0882285118}{0.0016561044}\right)^2 = 2838.196$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel X_3 & X_5 memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(7.0732874 + 0.0000307X_3 + 0.0882285X_5) + \varepsilon_i$$

j. Model dengan hanya X_4 & X_5 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.19 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.19 Taksiran parameter untuk model dengan hanya melibatkan X_4 & X_5 sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	11.37708627...	0.02070819...

X4	0.00034218...	0.00000159...
X5	-0.36925866...	0.00294727...

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_4 = \left(\frac{\hat{\beta}_4}{SE(\hat{\beta}_4)}\right)^2 = \left(\frac{0.00034218}{0.00000159}\right)^2 = 46238.133$$

$$W_5 = \left(\frac{\hat{\beta}_5}{SE(\hat{\beta}_5)}\right)^2 = \left(\frac{-0.36925866}{0.00294727}\right)^2 = 15697.130$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel X_4 & X_5 memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(11.377086 + 0.000342X_4 - 0.369258X_5) + \varepsilon_i$$

3. Model dengan melibatkan tiga variabel prediktor

   Berikut ini beberapa model yang akan terbentuk jika hanya melibatkan tiga variabel prediktor saja pada model. Hasil analisis selengkapnya untuk model dengan tiga variabel prediktor dapat dilihat pada lampiran 6.

   a. Model dengan hanya $X_1, X_2,$ & X_3 sebagai variabel prediktor

      Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.20 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.20 Taksiran parameter untuk model dengan hanya melibatkan $X_1, X_2, \& X_3$ sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	6.8621399971…	0.0215755990…
X1	0.0005356057…	0.0001192946…
X2	0.0696267148…	0.0032858326…
X3	0.0000271893…	0.0000001527…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_1 = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{0.0005356057}{0.0001192946}\right)^2 = 20.158$$

$$W_2 = \left(\frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}\right)^2 = \left(\frac{0.0696267148}{0.0032858326}\right)^2 = 449.015$$

$$W_3 = \left(\frac{\hat{\beta}_3}{SE(\hat{\beta}_3)}\right)^2 = \left(\frac{0.0000271893}{0.0000001527}\right)^2 = 31707.708$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel $X_1, X_2, \& X_3$ memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(6.8621399 + 0.0005356X_1 + 0.0696267X_2 + 0.0000271X_3) + \varepsilon_i$$

b. Model dengan hanya $X_1, X_2, \& X_4$ sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.21 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.21 Taksiran parameter untuk model dengan hanya melibatkan $X_1, X_2, \& X_4$ sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	6.7297506…	0.0227696…
X1	-0.0131781…	0.0001298…
X2	0.3397366…	0.0029415…
X4	0.0001795…	0.0000014…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_1 = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{-0.0131781}{0.0001298}\right)^2 = 10294.933$$

$$W_2 = \left(\frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}\right)^2 = \left(\frac{0.3397366}{0.0029415}\right)^2 = 13338.960$$

$$W_4 = \left(\frac{\hat{\beta}_4}{SE(\hat{\beta}_4)}\right)^2 = \left(\frac{0.0001795}{0.0000014}\right)^2 = 14497.107$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel $X_1, X_2, \& X_4$ memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(6.7297506 - 0.0131781X_1 + 0.3397366X_2 + 0.0001795X_4) + \varepsilon_i$$

c. Model dengan hanya $X_1, X_2, \& X_5$ sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.22 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.22 Taksiran parameter untuk model dengan hanya melibatkan $X_1, X_2, \& X_5$ sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	6.0412…	0.0265…
X1	-0.0046…	0.0001…
X2	0.3317…	0.0028…
X5	-0.0814…	0.0032…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_1 = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{-0.0046}{0.0001}\right)^2 = 1100.432$$

$$W_2 = \left(\frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}\right)^2 = \left(\frac{0.3317}{0.0028}\right)^2 = 14378.589$$

$$W_5 = \left(\frac{\hat{\beta}_5}{SE(\hat{\beta}_5)}\right)^2 = \left(\frac{-0.0814}{0.0032}\right)^2 = 658.012$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel $X_1, X_2, \& X_5$ memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(6.0412 - 0.0046X_1 + 0.3317X_2 - 0.0814X_5) + \varepsilon_i$$

d. Model dengan hanya $X_1, X_3, \& X_4$ sebagai variabel predicto

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.23 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.23 Taksiran parameter untuk model dengan hanya melibatkan $X_1, X_3, \& X_4$ sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	8.1235708...	0.0146072...
X1	-0.0017951...	0.0000794…
X3	0.0000264...	0.0000001...
X4	0.0001201...	0.0000016...

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_1 = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{-0.0017951}{0.0001}\right)^2 = 511.081$$

$$W_3 = \left(\frac{\hat{\beta}_3}{SE(\hat{\beta}_3)}\right)^2 = \left(\frac{0.0000794}{0.0000001}\right)^2 = 39759.161$$

$$W_4 = \left(\frac{\hat{\beta}_4}{SE(\hat{\beta}_4)}\right)^2 = \left(\frac{0.0001201}{0.0000016}\right)^2 = 5392.710$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel $X_1, X_3, \& X_4$ memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(8.1235708 - 0.0017951X_1 + 0.0000794X_3$$
$$+ 0.0001201X_4) + \varepsilon_i$$

e. Model dengan hanya $X_1, X_3, \& X_5$ sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.24 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.24 Taksiran parameter untuk model dengan hanya melibatkan $X_1, X_3, \& X_5$ sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	7.30520026…	0.01506280…
X1	0.00307996…	0.00008734…
X3	0.00002896…	0.00000013…
X5	-0.00910859…	0.00319823…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_1 = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{0.00307996}{0.00008734}\right)^2 = 1243.422$$

$$W_3 = \left(\frac{\hat{\beta}_3}{SE(\hat{\beta}_3)}\right)^2 = \left(\frac{0.00002896}{0.00000013}\right)^2 = 48632.516$$

$$W_5 = \left(\frac{\hat{\beta}_5}{SE(\hat{\beta}_5)}\right)^2 = \left(\frac{-0.00910859}{0.00319823}\right)^2 = 8.111$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel $X_1, X_3, \& X_5$ memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(7.30520026 - 0.00307996X_1 + 0.00002896X_3 - 0.00910859X_5) + \varepsilon_i$$

f. Model dengan hanya $X_1, X_4, \& X_5$ sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.25 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.25 Taksiran parameter untuk model dengan hanya melibatkan $X_1, X_4, \& X_5$ sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	11.1854749...	0.0203532...
X1	0.0041619...	0.0000864…
X4	0.0003027…	0.0000017...
X5	-0.4406952…	0.0032560…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_1 = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{0.0041619}{0.0000864}\right)^2 = 2317.003$$

$$W_4 = \left(\frac{\hat{\beta}_4}{SE(\hat{\beta}_4)}\right)^2 = \left(\frac{0.0003027}{0.0000017}\right)^2 = 30691.918$$

$$W_5 = \left(\frac{\hat{\beta}_5}{SE(\hat{\beta}_5)}\right)^2 = \left(\frac{-0.4406952}{0.0032560}\right)^2 = 18318.514$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel $X_1, X_4, \& X_5$ memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(11.1854749 + 0.0041619X_1 + 0.0003027X_4 - 0.4406952X_5) + \varepsilon_i$$

g. Model dengan hanya $X_2, X_3, \& X_4$ sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.26 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.26 Taksiran parameter untuk model dengan hanya melibatkan $X_2, X_3, \& X_4$ sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	7.9420796174…	0.0228713789…
X2	-0.0107174678…	0.0019071079…
X3	0.0000267938…	0.0000001412…
X4	0.0000953999…	0.0000014145…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_2 = \left(\frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}\right)^2 = \left(\frac{-0.0107174678}{0.0019071079}\right)^2 = 31.581$$

$$W_3 = \left(\frac{\hat{\beta}_3}{SE(\hat{\beta}_3)}\right)^2 = \left(\frac{0.0000267938}{0.0000001412}\right)^2 = 36017.737$$

$$W_4 = \left(\frac{\hat{\beta}_4}{SE(\hat{\beta}_4)}\right)^2 = \left(\frac{0.0000953999}{0.0000014145}\right)^2 = 4548.448$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel $X_2, X_3, \& X_4$memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(7.9420796 - 0.0107174X_2 + 0.0000267X_3 + 0.0000953X_4) + \varepsilon_i$$

h. Model dengan hanya $X_2, X_3, \& X_5$ sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.27 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.27 Taksiran parameter untuk model dengan hanya melibatkan $X_2, X_3, \& X_5$ sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	6.7873505253…	0.0154877051…
X2	0.0794907388…	0.0019321122…
X3	0.0000270479…	0.0000001506…
X5	0.0066253423…	0.0025888881…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_2 = \left(\frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}\right)^2 = \left(\frac{0.0794907388}{0.0019321122}\right)^2 = 1692.654$$

$$W_3 = \left(\frac{\hat{\beta}_3}{SE(\hat{\beta}_3)}\right)^2 = \left(\frac{0.0000270479}{0.0000001506}\right)^2 = 32269.582$$

$$W_5 = \left(\frac{\hat{\beta}_5}{SE(\hat{\beta}_5)}\right)^2 = \left(\frac{0.0066253423}{0.0025888881}\right)^2 = 6.549$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel $X_2, X_3, \& X_5$memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(6.7873505 + 0.0794907X_2 + 0.0000270X_3$$
$$+ 0.0066253X_5) + \varepsilon_i$$

i. Model dengan hanya $X_2, X_4, \& X_5$ sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.28 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.28 Taksiran parameter untuk model dengan hanya melibatkan $X_2, X_4, \& X_5$ sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	9.95380431…	0.02540083…
X2	0.17016670…	0.00172417…
X4	0.00027453…	0.00000172…
X5	-0.46789276…	0.00328034…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_2 = \left(\frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}\right)^2 = \left(\frac{0.17016670}{0.00172417}\right)^2 = 9740.663$$

$$W_4 = \left(\frac{\hat{\beta}_4}{SE(\hat{\beta}_4)}\right)^2 = \left(\frac{0.00027453}{0.00000172}\right)^2 = 25396.600$$

$$W_5 = \left(\frac{\hat{\beta}_5}{SE(\hat{\beta}_5)}\right)^2 = \left(\frac{-0.46789276}{0.00328034}\right)^2 = 20344.92587$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel $X_2, X_4, \& X_5$memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(9.95380431 + 0.17016670X_2 + 0.00027453X_4 - 0.46789276X_5) + \varepsilon_i$$

j. Model dengan hanya $X_3, X_4, \& X_5$ sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.29 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.29 Taksiran parameter untuk model dengan hanya melibatkan $X_3, X_4, \& X_5$ sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	9.3248338926…	0.0260172366…
X3	0.0000229168…	0.0000001459…
X4	0.0001877588…	0.0000019107…
X5	-0.1940362284…	0.0033117317…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_3 = \left(\frac{\hat{\beta}_3}{SE(\hat{\beta}_3)}\right)^2 = \left(\frac{0.0000229168}{0.0000001459}\right)^2 = 24667.270$$

$$W_4 = \left(\frac{\hat{\beta}_4}{SE(\hat{\beta}_4)}\right)^2 = \left(\frac{0.0001877588}{0.0000019107}\right)^2 = 9656.325$$

$$W_5 = \left(\frac{\hat{\beta}_5}{SE(\hat{\beta}_5)}\right)^2 = \left(\frac{-0.1940362284}{0.0033117317}\right)^2 = 3432.854$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel $X_3, X_4, \& X_5$memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(9.3248338 + 0.0000229X_3 + 0.0001877X_4 - 0.1940362X_5) + \varepsilon_i$$

4. Model dengan melibatkan empat variabel prediktor

Berikut ini beberapa model yang akan terbentuk jika hanya melibatkan empat variabel prediktor saja pada model. Hasil analisis selengkapnya untuk model dengan tiga variabel prediktor dapat dilihat pada lampiran 7.

a. Model dengan hanya X_1, X_2, X_3 & X_4 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.30 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.30 Taksiran parameter untuk model dengan hanya melibatkan X_1, X_2, X_3 & X_4 sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	7.63942776...	0.02483718...
X1	-0.00467495...	0.00014574...
X2	0.08378730...	0.00349914...
X3	0.00002426...	0.00000016...
X4	0.00012148...	0.00000163...

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_1 = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{-0.00467495}{0.00014574}\right)^2 = 1028.872$$

$$W_2 = \left(\frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}\right)^2 = \left(\frac{0.08378730}{0.00349914}\right)^2 = 573.366$$

$$W_3 = \left(\frac{\hat{\beta}_3}{SE(\hat{\beta}_3)}\right)^2 = \left(\frac{0.00002426}{0.00000016}\right)^2 = 23009.733$$

$$W_4 = \left(\frac{\hat{\beta}_4}{SE(\hat{\beta}_4)}\right)^2 = \left(\frac{0.00012148}{0.00000163}\right)^2 = 5522.251$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel X_1, X_2, X_3 & X_4memiliki pengaruh secara nyata terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(7.63942776 - 0.00467495X_1 + 0.08378730X_2 + 0.00002426X_3 + 0.00012148X_4) + \varepsilon_i$$

b. Model dengan hanya X_1, X_2, X_3 & X_5 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.31 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.31 Taksiran parameter untuk model dengan hanya melibatkan X_1, X_2, X_3 & X_5 sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	6.86482881…	0.02605696...
X1	0.00055224...	0.00014968...
X2	0.06954502...	0.00331541...
X3	0.00002718...	0.00000015...
X5	-0.00059667...	0.00324306...

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_1 = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{0.00055224}{0.00014968.}\right)^2 = 13.610$$

$$W_2 = \left(\frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}\right)^2 = \left(\frac{0.06954502}{0.00331541}\right)^2 = 440.003$$

$$W_3 = \left(\frac{\hat{\beta}_3}{SE(\hat{\beta}_3)}\right)^2 = \left(\frac{0.00002718}{0.00000015}\right)^2 = 30674.238$$

$$W_5 = \left(\frac{\hat{\beta}_5}{SE(\hat{\beta}_5)}\right)^2 = \left(\frac{-0.00059667}{0.00324306}\right)^2 = 0.0338$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel $X_1, X_2, \& X_3$ memiliki pengaruh secara nyata terhadap variabel respon dan X_5 dikatakan tidak memiliki pengaruh terhadap variabel respon. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(6.86482881 + 0.00055224X_1 + 0.06954502X_2 - 0.00002718X_3) + \varepsilon_i$$

c. Model dengan hanya $X_1, X_2, X_4 \& X_5$ sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.32 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.32 Taksiran parameter untuk model dengan hanya melibatkan $X_1, X_2, X_4 \& X_5$ sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	9.3384965...	0.0300937…
X1	-0.0067019…	0.0001488...
X2	0.2809067...	0.0030700…
X4	0.0002940...	0.0000018...
X5	-0.4175052...	0.0036378…

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_1 = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{-0.0067019}{0.0001488}\right)^2 = 2028.057$$

$$W_2 = \left(\frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}\right)^2 = \left(\frac{0.2809067}{0.0030700}\right)^2 = 8371.956$$

$$W_4 = \left(\frac{\hat{\beta}_4}{SE(\hat{\beta}_4)}\right)^2 = \left(\frac{0.0002940}{0.0000018}\right)^2 = 24881.803$$

$$W_5 = \left(\frac{\hat{\beta}_5}{SE(\hat{\beta}_5)}\right)^2 = \left(\frac{-0.4175052}{0.0036378}\right)^2 = 13171.611$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel $X_1, X_2, X_4, \& X_5$. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(9.3384965 - 0.0067019X_1 + 0.2809067X_2 + 0.0002940X_4 - 0.4175052X_5) + \varepsilon_i$$

d. Model dengan hanya X_1, X_3, X_4 & X_5 sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.33 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.33 Taksiran parameter untuk model dengan hanya melibatkan X_1, X_3, X_4 & X_5 sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	9.31618060...	0.02589060...
X1	0.00085730...	0.00009390...
X3	0.00002259...	0.00000015...
X4	0.00018197...	0.00000200...
X5	-0.21151411...	0.00381237...

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_1 = \left(\frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}\right)^2 = \left(\frac{0.00085730}{0.00009390}\right)^2 = 83.341$$

$$W_3 = \left(\frac{\hat{\beta}_3}{SE(\hat{\beta}_3)}\right)^2 = \left(\frac{0.00002259}{0.00000015}\right)^2 = 22535.073$$

$$W_4 = \left(\frac{\hat{\beta}_4}{SE(\hat{\beta}_4)}\right)^2 = \left(\frac{0.00018197}{0.00000200}\right)^2 = 8260.317$$

$$W_5 = \left(\frac{\hat{\beta}_5}{SE(\hat{\beta}_5)}\right)^2 = \left(\frac{-0.21151411}{0.00381237}\right)^2 = 3078.136$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel $X_1, X_3, X_4, \& X_5$. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(9.31618060 + 0.00085730X_1 + 0.00002259X_3 + 0.00018197X_4 - 0.21151411X_5) + \varepsilon_i$$

e. Model dengan hanya $X_2, X_3, X_4 \;\&\; X_5$ sebagai variabel prediktor

Hasil penaksiran parameter $\boldsymbol{\beta}$ diperlihatkan pada Tabel 4.34 diikuti dengan pengujian parameter dan model yang terbentuk.

Tabel 4.34 Taksiran parameter untuk model dengan hanya melibatkan $X_2, X_3, X_4 \;\&\; X_5$ sebagai variabel prediktor

Parameter	β	Std. Error
(Intercept)	9.06902602...	0.02843942...
X2	0.04750817...	0.00211167...
X3	0.00002099...	0.00000017...

X4	0.00017979...	0.00000194...
X5	-0.23197372...	0.00375789...

Selanjutnya dilakukan pengujian parameter menggunakan uji Wald,

$$W_2 = \left(\frac{\hat{\beta}_2}{SE(\hat{\beta}_2)}\right)^2 = \left(\frac{0.04750817}{0.00211167}\right)^2 = 506.154$$

$$W_3 = \left(\frac{\hat{\beta}_3}{SE(\hat{\beta}_3)}\right)^2 = \left(\frac{0.00002099}{0.00000017}\right)^2 = 15246.922$$

$$W_4 = \left(\frac{\hat{\beta}_4}{SE(\hat{\beta}_4)}\right)^2 = \left(\frac{0.00017979}{0.00000194}\right)^2 = 8530.868$$

$$W_5 = \left(\frac{\hat{\beta}_5}{SE(\hat{\beta}_5)}\right)^2 = \left(\frac{-0.23197372}{0.00375789}\right)^2 = 3810.561$$

berdasarkan uji parameter yang telah dilakukan dengan kriteria uji Wald maka dapat dijelaskan variabel $X_2, X_3, X_4, \& X_5$. Sehingga didapat model sebagai berikut:

$$\hat{y} = \exp(9.06902602 + 0.04750817X_2 + 0.00002099X_3 + 0.00017979X_4 - 0.23197372X_5) + \varepsilon_i$$

B. Pembahasan

a. Pemilihan model terbaik

Setelah didapatkan beberapa model dengan metode regresi Poisson biasa dan *Generalized Poisson Regression* maka dilakukan pemilihan model terbaik berdasarkan kriteria nilai AIC.

1. Model terbaik dengan melibatkan satu variabel prediktor

Tabel 4.35 Nilai AIC untuk setiap model dengan hanya satu variabel prediktor

Model	Nilai AIC
X1	78242.575
X2	61936.116
X3	29380.390
X4	57744.933
X5	89042.062

Berdasarkan Tabel 4.32 maka dapat dijelaskan bahwa model terbaik dengan satu variabel prediktor adalah model yang hanya melibatkan variabel X_3 pada model, dipilih berdasarkan kriteria nilai AIC yang terkecil.

2. Model terbaik dengan melibatkan dua variabel prediktor

Tabel 4.36 Nilai AIC untuk setiap model dengan hanya dua variabel prediktor

Model	Nilai AIC
X1, X2	56999.941
X1, X3	25451.813
X1, X4	57057.721
X1, X5	73159.661
X2, X3	25019.546
X2, X4	54250.247
X2, X5	57478.515
X3, X4	20425.680
X3, X5	26681.426
X4, X5	40599.426

Berdasarkan Tabel 4.33 maka dapat dijelaskan bahwa model terbaik dengan dua variabel prediktor adalah model yang hanya melibatkan variabel X_3 & X_4 pada model, dipilih berdasarkan kriteria nilai AIC yang terkecil.

3. Model terbaik dengan melibatkan tiga variabel prediktor

Tabel 4.37 Nilai AIC untuk setiap model dengan hanya tiga variabel prediktor

Model	Nilai AIC
X1, X2, X3	25001.459
X1, X2, X4	41771.553
X1, X2, X5	56347.418
X1, X3, X4	19906.105
X1, X3, X5	25445.707
X1, X4, X5	38304.742
X2, X3, X4	20395.957
X2, X3, X5	25015.000
X2, X4, X5	31045.703
X3, X4, X5	16970.179

Berdasarkan Tabel 4.34 maka dapat dijelaskan bahwa model terbaik dengan dua variabel prediktor adalah model yang hanya melibatkan variabel $X_3, X_4,$ & X_5 pada model, dipilih berdasarkan kriteria nilai AIC yang terkecil.

4. Model terbaik dengan melibatkan empat variabel prediktor

Tabel 4.38 Nilai AIC untuk setiap model dengan hanya empat variabel prediktor

Model	Nilai AIC
X1, X2, X3, X4	19328.916
X1, X2, X3, X5	25003.425
X1, X2, X4, X5	28937.663
X1, X3, X4, X5	16889.102
X2, X3, X4, X5	16473.095

Berdasarkan Tabel 4.35 maka dapat dijelaskan bahwa model terbaik dengan dua variabel prediktor adalah model yang hanya melibatkan variabel $X_2, X_3, X_4, \& X_5$ pada model, dipilih berdasarkan kriteria nilai AIC yang terkecil.

5. Model terbaik dari semua model yang terbentuk.

Tabel 4.39 Nilai AIC dari semua model terbaik yang diplih

Model	Nilai AIC
X3	29380.390
X3, X4	20425.680
X3, X4, X5	16970.179
X2, X3, X4, X5	16473.095
X1, X2, X3, X4, X5	16277.680

Berdasarkan Tabel 4.36 maka dapat dijelaskan bahwa model terbaik berdasarkan kriteria nilai AIC adalah model yang melibatkan semua variabel prediktor pada model, namun karena data yang diamati mengalami kasus Overdispersi sehingga pemodelan angka putus sekolah di Sulawesi Selatan menggunakan model regresi Poisson dikatakan tidak layak digunakan. Oleh karena itu, model terbaik untuk pemodelan angka putus sekolah di Provinsi Sulawesi Selatan adalah menggunakan *Generalized Poisson Regression* dengan model sebagai berikut

$$y = \exp(9.06902602 + 0.04750817X_2 + 0.00002099X_3 + 0.00017979X_4 - 0.23197372X_5) + \varepsilon_i$$

b. Interpretasi Hasil

Setelah mendapatkan model terbaik dari beberapa model yang terbentuk maka selanjutnya akan dilakukakan interpretasi model.

$$\hat{y} = \exp(9.06902602 + 0.04750817X_2 + 0.00002099X_3 + 0.00017979X_4 - 0.23197372X_5) + \varepsilon_i$$

Berdasarkan model yang diperoleh maka dapat dikatakan bahwa variabel yang berpengaruh secara nyata terhadap model adalah, rasio siswa terhadap guru, jumlah penduduk miskin, kepadatan penduduk, dan rata-rata lama sekolah. Pada model dapat dilihat bahwa angka putus sekolah bagi anak usia wajib belajar di Provinsi Sulawesi Selatan akan bertambah sebesar 0.04750817 jika rasio siswa terhadap sekolah bertambah satu satuan, akan bertambah sebesar 0.00002099 jika penduduk miskin bertambah sebesar satu satuan, dan juga akan bertambah sebesar 0.00017979 jika kepadatan penduduk bertambah sebesar satu satuan.

Sebaliknya, angka putus sekolah bagi anak usia wajib belajar akan berkurang sebesar 0.23197372 jika rata-rata lama bersekolah bertambah sebesar satu satuan, ngka putus sekolah usia wajib belajar juga akan bertambah sebesar 0.0001877907 jika kepadatan penduduk bertambah sebesar satu satuan. Dengan demikian, dapat diartikan bahwa rata-rata angka putus sekolah bagi usia wajib belajar di Provinsi Sulawesi Selatan dapat berkurang jika rata-rata lama sekolah bertambah.

BAB V

KESIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan maka diperoleh kesimpulan:

1. Penaksiran parameter model *Generalized Poisson Regression* pada dasarnya sama dengan penaksiran parameter model regresi Poisson, yaitu:
   a. Fungsi likelihood model GPR

   $L(\boldsymbol{\beta}, k, y) =$

   $$\prod_{i=1}^{n} \left(\frac{\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}\right)}{1 + k \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}\right)} \right)^{y_i} \frac{(1 + k y_i)^{y_i - 1}}{y_i!} \exp\left(- \frac{\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}\right)(1 + k y_i)}{\left(1 + k \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}\right)\right)} \right) \Bigg\}$$

   b. Fungsi log-likelihood model GPR

   $$\ln L(\boldsymbol{\beta}, k, y) = \sum_{i=1}^{n} \{ y_i (x_i^T \beta) - y_i \ln(1 + k \exp(x_i^T \beta)) + (y_i - 1) \ln(1 + k y_i) - \frac{\exp(x_i^T \beta)(1 + k y_i)}{(1 + k \exp(x_i^T \beta))} - \ln y_i! \}$$

   c. Hasil penurunan persamaan log-likelihood

   - $\frac{\partial \ln L(\beta, k)}{\partial \beta_0} = 0$

     $\sum_{i=1}^{n} \left(\frac{y_i - \mu_i}{(1 + k \mu_i)^2} \right) = 0$

   - $\frac{\partial \ln L(\beta, k)}{\partial \beta_j} = 0$

     $\sum_{i=1}^{n} \frac{x_{ij}(y_i - \mu_i)}{(1 + k \mu_i)^2} = 0$

   - $\frac{\partial \ln L(\beta, k)}{\partial k} = 0$

     $\sum_{i=1}^{n} \left\{ - \frac{y_i \mu_i}{1 + k \mu_i} + \frac{y_i (y_i - 1)}{1 + k y_i} - \frac{\mu_i (y_i - \mu_i)}{(1 + k \mu_i)^2} \right\} = 0$

2. Model angka putus sekolah usia wajib belajar di provinsi Sulawesi selatan menggunakan pendekatan Generalized Poisson Regression, yaitu:

$$\hat{y} = \exp(9.06902602 + 0.04750817X_2 + 0.00002099X_3 + 0.00017979X_4 - 0.23197372X_5) + \varepsilon_i$$

3. Faktor-faktor yang berpengaruh secara signifikan terhadap pertumbuhan jumlah angka putus sekolah di Provinsi Sulawesi Selatan berdasarkan model Generalized Poisson Regression, yaitu rasio siswa terhadap guru, jumlah penduduk miskin, kepadatan penduduk, dan rata-rata lama

B. Saran

Adapun saran yang dapat diberikan,yaitu:

1. Perhitungan estimasi parameter dalam penelitian ini sepenuhnya dilakukan menggunakan software SPSS 23, penelitian selanjutnya dengan melakukan perhitungan manual menggunakan Microsoft Excel sangat diharapkan.
2. Diharapkan penelitian selanjutnya mengenai pemodelan angka putus sekolah di Sulawesi Selatan dapat dilakukan dengan memperhitungkan faktor letak geografis.

DAFTAR PUSTAKA

Agresti, A. (2002). *Categorical Data Analysis*. New York: John Wiley & Sons Inc.

Astari, G. R., & Srinadi, G. M. (2013). Pemodelan Jumlah Anak Putus Sekolah di Provinsi Bali dengan Pendekatan Semi-Parametric Geographically Weighted Poisson Regression. *E-Jurnal Matematika Universitas Udayana Bali*, Vol.2 No.3, 29-34.

Astuti, C. C., Sumarminigsih, E., Soehono, L. A. (2013). Perbandingan Generalized Poisson Regression dan Negative Binomial Negative untuk Data Overdispersi dan Underdispersi pada Regresi Poisson. *Jurnal Mahasiswa Matematika Universitas Brawijaya Malang*, Vol.1 No.2, 105-108.

Aulele, S. N. (2012). Pemodelan Jumlah Kematian Bayi di Provinsi Maluku Tahun 2010 dengan Menggunakan Regresi Poisson. *Jurnal Barekeng Universitas Pattimura Ambon*, Vol.5 No.2, 23-27.

Badan Pusat Statistik. (2015). *Statistik Daerah Provinsi Sulawesi Selatan 2015*. Makassar: Badan Pusat Statistik Kota Makassar.

Cahyandari, R. (2014). Pengujian Overdispersi pada Model Regresi Poisson. *Statisika Universitas Islam Bandung*, Vol.14 No.2, 69-76.

Fadhillah, F. (2011). Aplikasi Regresi Binomial Negatif dan Generalized Poisson Regression dalam Mengatasi Overdispersi pada Regresi Poisson. *Skripsi*. Fakultas Sains dan Teknologi Universitas Islam Negeri Syarif Hidayatullah. Jakarta.

Fadmi, F. R. (2014). Pemodelan Geographically Weighted Poisson Regression (GWPR pada Jumlah Kasus Baru Penyakit Kusta di Kabupaten Buton Provinsi Sulawesi Tenggara Tahun 2013. *Tesis*. Fakultas Kesehatan Masyarakat Universitas Airlangga. Surabaya.

Fitroni, B. N., & Zain, I. (2013). Pemodelan Angka Putus Sekolah Usia Wajib Belajar Menggunakan Metode Regresi Spasial di Jawa Timur. *Jurnal Sains dan Seni Pomits Institut Teknologi Sepuluh Nopember*, Vol.2 No.2, 171-176.

Irwan, & Sari, D. P. (2013). Pemodelan Regresi Poisson, Binomial Negatif dan pada Kasus Kecelakaan Kendaraan Bermotor di Lalu Lintas Sumatera Barat. *Seminar Nasional Matematika dan Pendidikan Matematika FMIPA UNY*, 107-122. Universitas Negeri Yogyakarta.

Krisnawardhani, T., Salam, N., & Anggraini, D. (2010). Analisis Regresi Linear Berganda dengan Satu Variabel Boneka (Dummy Variable). *Jurnal Matematika Murni dan Terapan Epsilon Universitas Lambung Mangkurat*, Vol.4 No. 2, 14-20

Myers, R. H. (1990). *Classical and Modern Regression with Application* (2nd ed.). Boston: PWS-KENT Publishing Company.

Pratiwi, D. N. (2012). Mengatasi Overdispersi pada Regresi Poisson dengan Regresi Generalized Poisson. *Skripsi*. Fakultas MIPA Universitas Pendidikan Indonesia. Bandung.

Putra, I. Y. E., Kencana, I. E., & Srinadi, I. A. (2013). Penerapan Regresi Generalized Poisson untuk Mengatasi Fenomena Overdispersi pada Kasus Regresi Poisson. *E-Jurnal Matematika Universitas Udayana Bali*, Vol.2 No.2, 49-53.

Riyadi, Hartini, S., dkk. (2015). *Indikator Kesejahteraan Rakyat 2015*. Jakarta: Badan Pusat Statistik (BPS).

Rohma, A. (2013). Perbandingan Metode Ridge dengan Metoda PCA (Principal Component Analysis) dalam Mengatasi Multikolinieritas. *Skripsi*. Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga. Yogyakarta.

Ruliana. (2015). Pemodelan Generalized Poisson Regression (GPR) untuk Mengatasi Pelanggaran Equidispersi pada Regresi Poisson Kasus Campak di Kota Semarang. *Skripsi*. Fakultas MIPA Universitas Negeri Semarang. Semarang.

Sembiring, R. K. (1995). *Analisis Regresi*. Bandung: ITB Bandung.

Simarmata, T. R., & Ispriyanti, D. (2011). Penanganan Overdispersi pada Model Regresi Poisson Menggunakan Model Regresi Binomial Negatif. *Media Statiska Universitas Diponegoro*, Vol.2 No.2, 95-104.

Syafiee, H. I. (2005). *Pengantar Ilmu Pengetahuan*. Bandung: PT Refika Aditama.

Walpole, E. R., & Myers, H. R. (1995). *Ilmu Peluang dan Statistik untuk Insinyur dan Ilmuwan*. Bandung: ITB.

Yulianingsih, K. A., Sukarsa, K. G., & Suciptawati, L. (2012). Penerapan Regresi Poisson untuk Mengetahui Faktor-faktor yang Mempengaruhi Jumlah Siswa SMA/SMK yang Tidak Lulus UN di Bali. *E-Jurnal Matematika Universitas Udayana*, Vol.1 No.1, 59-63